Indian Statistical Institute

Semester-1 2012-2013

M.Tech.(CS) - Second Year

Mid-semester Examination (15 September, 2012)

Subject: Compiler Construction

Maximum marks: 40 Total marks: 45 Duration 2.5 hrs.

Please keep your answers brief and to the point.

1. Tables in HTML^{--}, a hypothetical markup language, are enclosed in `<table>` and `</table>` tags. Further,

   - each table consists of one or more rows enclosed in `<tr>` and `</tr>` tags;
   - each row consists of one or more columns (or cells) enclosed in `<td>` and `</td>` tags;
   - the number of columns in different rows may be different;
   - no nesting of tables is permitted in HTML^{--};
   - tags do not have attributes in HTML^{--} (specifically, the `colspan` and `rowspan` attributes cannot be used).

   Write a (single) lex program that takes an HTML^{--} table as input and

   - checks that the tags in the given table are properly nested (i.e., tables are not nested, rows occur only within tables, but never within `<tr>` or `<td>` tags, etc.),
   - computes the total number of rows in the table,
   - computes the maximum number of columns across all rows.

   You may assume that the only tags present in the input are the ones listed above. [12]

   Note 1: You may make reasonable assumptions where necessary. **Clearly state any assumptions you make.**

   Note 2: In Lex, . is a regular expression that matches any single character other than the newline.

   Warning: Lex always tries to find the longest possible match for any given pattern.

2. Let G be a grammar consisting of the following productions. $S, L, E, Elist$ are non-terminals (S being the start symbol), while all other symbols are terminals.

$$S \rightarrow L = E \qquad E \rightarrow L \qquad L \rightarrow \mathbf{id} \qquad L \rightarrow Elist\ ]$$

$$Elist \rightarrow \mathbf{id}\ [\ E \qquad Elist \rightarrow Elist, E$$

   (a) Eliminate left-recursion from the above grammar.

   (b) Compute the *FIRST* set for the right hand side of each production of the modified (non-left-recursive) grammar.

   (c) Compute the *FOLLOW* set for each non-terminal of the modified grammar. Show your rough work.

(d) Construct the $LL(1)$ parsing table for the modified grammar. Show that the modi grammar is not $LL(1)$.

[2+3+4+4=

3. Consider the following grammar (capital letters denote non-terminals, small letters denote t minals) with **10** productions.

$$S \to aAa \quad S \to aBb \quad S \to cCc \quad S \to dDd$$

$$A \to Ax \quad A \to z \quad B \to By \quad B \to z$$

$$C \to Ac \quad D \to Bc$$

(a) Construct the canonical collection of $LR(1)$ (i.e. *canonical LR*) items for this grammar. You should get about 24 sets in your answer.

(b) How many states does the LALR parsing table for this grammar contain? Justify yo answer.

(c) Calculate the *FOLLOW* sets for A and B. Hence show that the gramar is not an SL grammar. You do not need to construct the complete SLR parsing table.

[12+3+(2+3)=2

INDIAN STATISTICAL INSTITUTE

Mid Semestral Examination:(2012-2013)

MTech C.S. 2nd Year

Digital Signal Processing

Date: 17.9.2012 Maximum Marks: 60 Duration: 2 hours

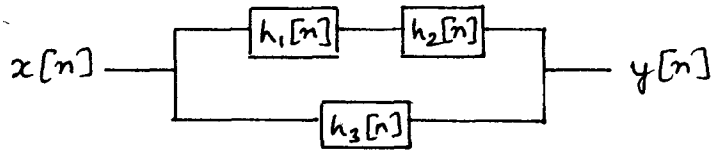
Note: The marks add up to 71. The maximum you can score is 60. The exam is open-book, open-notes. Use of calculators is permitted.

1. For an input $x[n]$, a discrete-time system has the output

$$y[n] = 2 + x[3 - n]$$

Determine and justify if the system is linear, stable, causal and time-invariant. [10]

2. Three stable LTI (linear, time-invariant) systems are connected to form the discrete-time system shown below.



Compute the overall impulse response for

$$h_1[n] = \delta[n-2]+\delta[n+1], \quad h_2[n] = \delta[n-1]+2\delta[n+2], \quad h_2[n] = \delta[n-5]+\delta[n]$$

[5]

3. Can the sequence $a^n u[n]$, where a is complex and $u[n]$ is the unit step sequence, be the eigenfunction of an LTI system? [5]

4. Determine the z-transforms of

   (a) $x[n] = 0.6^{|n|}$. Indicate the ROC (region of convergence).

   (b) $n^2 x[n]$ in terms of $X(z)$, the z-transform of $x[n]$ and its ROC R_x.

[5+5]

5. Determine the inverse z-transforms of

   (a) $z^{-2}X(z^6)$ without actually computing $X(z)$, where $x[n] = 0.4^n u[n]$

   (b)

$$\frac{1}{z(z-1/4)(z-1/2)(z-4/3)}$$

   such that $x[n]$ is stable. Evaluate its value at $n = 3$ using contour integration.

[8+8]

6. Let $X(e^{j\omega})$ denote the Discrete-time Fourier Transform (DTFT) of the sequence

$$x[n] = \{3,\ 1,\ -5,\ 2,\ 4\} \quad -2 \le n \le 2$$

   Without actually computing the DTFT, evaluate

   (a) $X(e^{j0})$

   (b) $X(e^{j\pi})$

   (c) $\int_{-\pi}^{\pi} X(e^{j\omega})d\omega$

   (d) $\int_{-\pi}^{\pi} |\frac{dX(e^{j\omega})}{d\omega}|^2 d\omega$

[10]

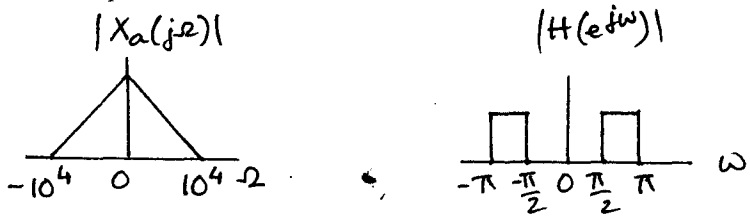
7. An LTI system has the difference equation

$$y[n] = 0.4y[n-1] + 0.05y[n-2] + 3x[n]$$

   (a) Determine the system function $H(z)$. Can the system be both causal and stable?

   (b) Determine the impulse response $h[n]$. [6+4]

8. An analog signal with Continuous Time Fourier Transform (CTFT) $X_a(j\Omega)$ as shown below is sampled at 20 kHz . It then passes through a filter with frequency response $H(e^{j\omega})$ as shown . Sketch the DTFT of the output. [5]



INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination: 2012 – 13

Course Name: M. Tech. (CS) - II

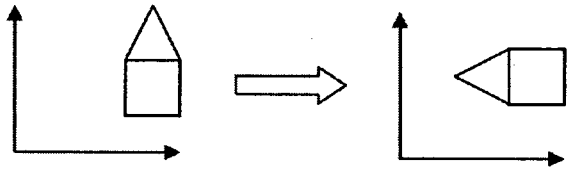
Subject: Computer Graphics

Date: 17. 09. 2012

Maximum Marks: 60 Duration: 2 hrs.

1. a) Why do we use a 4x4 homogenous coordinate transformation matrix when describing translations, rotations, and scales, etc.?

b) Provide a 3x3 matrix that will compute the new vertices of a planar house centered at **[10, 5]** after a rotation of 90 degrees about its center:

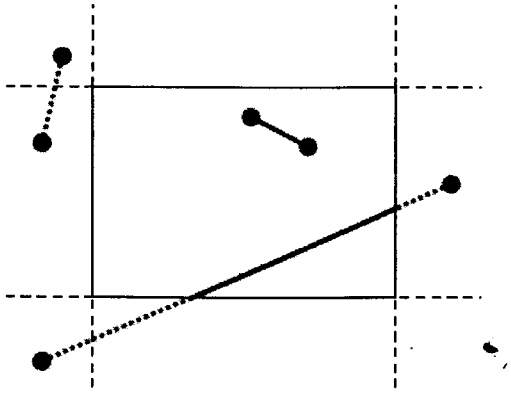


c) Calculate a 4 × 4 matrix to rotate points about the vector A = [1 1 1] by 90 degrees.

4+6+5 = 15

2. a) Suppose an RGB raster system is to be design using an 8-inch by 10-inch screen with a resolution of 100 pixels per inch in each direction. If we want to store 6 bits per pixel in the frame buffer, how much storage (in bytes) do we need for system the frame buffer?

b) Describe an algorithm for clipping a line against a rectangle.



Show that it works by using the above three examples. 5+(7+3) = 15

3. a) Write a simple function for scan conversion of lines with negative slope. Explain your steps.

b) Develop a scan conversion algorithm to draw an ellipse whose major axis is oriented at an angle specified by the user. 7+8=15

4. a) When using the Cohen-Sutherland line clipping algorithm, how do we check the outcodes to see if a line can be trivially accepted or rejected?

b) Develop and implement an algorithm for clipping a line against a triangle shaped window.

c) For scan line filling algorithm, what special treatments are to be done for horizontal and vertical edges of the polygon?

4+7+4=15

5. a) Consider the line which starts at (5, 14) and ends at (21, 2).

i) How many pixels will there be in this line?

ii) With the DDA (Digital Differential Analyzer) algorithm, what will be the amount added to the secondary component each time through the loop (incrementing value)?

b) Consider the two points A(1, 1) and B(4, 7). Let P(u)=(x(u),y(u)) be a parametric line function. What is a parametric equation of the line segment joining the points A and B?

c) "S" is a scaling that contracts everything by a factor of 1/3 along the x-axis, expands everything by a factor of 7 along the y axis, and leaves things unchanged along the z-axis. What is the 4x4 matrix representation for S?

d) If you rotate the point (20,30) by 90 degrees anticlockwise and then translate it by (-20,0) and then scale it by (2,1), where will the point be?

(2+2)+4+3+4 = 15

6. a) What are the window to viewport mappings for W = (0, 100, 50, 200) and V = (10, 50, 0, 60)?

b) Suppose a B´ezier curve C(u) is defined by the following four control points in the xy-plane: P0 =(−2, 0), P1 = (−2, 4), P2 = (2, 4) and P3 = (2, 0).

i) What is the degree of C(u)?

ii) What are the new control points if the degree of the original curve is increased by one?

c) Given a camera position ***P***, a vector normal to the image plane ***N***, and an up vector ***Vup***, describe how to convert a point W in world coordinates to a point in camera coordinates. Provide your final answer in the form of one (or a product of many) transformation matrix. Hint, the origin in camera coordinates is located at ***P*** and the world coordinate axes must be rotated to align with the camera's coordinate axes.

3+(2+4)+6 = 15

Indian Statistical Institute

M. TECH. (CS) 2 Year : 2012–2013
Mid-semester Examination
Subject: Cryptology

Date: 17/09/2012 Time: 2 hours Marks: 40

1. Prove that $|\mathsf{Prob}(l_u = f) - \frac{1}{2}| = \frac{|W_f(u)|}{2^{n+1}}$, where symbols have their usual meaning. [8]

2. Suppose a cryptosystem achieves perfect secrecy for a particular plaintext probability distribution. Prove that perfect secrecy is maintained for any plaintext probability distribution. [8]

3. Describe Finney state. Prove that an RC4 state is in Finney state if and only if it's prior state is in Finney state. [8]

4. Suppose a sequence of plaintext blocks, $x_1, \cdots, x_n$ yields the ciphertext sequence $y_1, \cdots, y_n$. Suppose that one ciphertext block, say y_i, is transmitted incorrectly. Show that number of blocks that will be decrypted incorrectly is equal to one if ECB or OFB modes are used for encryption; and equal to two if CBC or CFB modes are used. [8]

5. Give an algorithmic level description of AES. [8]

INDIAN STATISTICAL INSTITUTE

Mid Semestral Examination

M. Tech (CS) - II Year, 2012-2013 (Semester - III)

Advanced Algorithms for Graphs and Combinatorial Optimization Problems

Date: 19.09.2012 Maximum Marks : 50 Duration : 2.0 Hours

Note: The question paper is of 59 marks. You can answer as much as you can, but the maximum you can score is 50.

This is a 1 page question paper with 4 questions.

(Q1) Let $G = (V, E)$ be a weighted, directed graph with exactly one negative-weight edge and no negative-weight cycles. Give an algorithm to find the shortest distance from a s to all other vertices in V that has the same running time as Dijkstra. [6]

(Q2) Write down the MPM algorithm for finding a maximum flow. Prove the correctness of the algorithm. Derive worst case time complexity.

[8 + 12 + 10 = 30]

(Q3) Let $G = (V, E)$ be a bipartite graph with vertex partition $V = L \cup R$, and let $\mathcal{N}$ be its corresponding flow network. Give a good upper bound on the length of any augmenting path found in $\mathcal{N}$ during the execution of Ford-Fulkerson algorithm. [12]

(Q4) Bottleneck problem: Given a digraph G, $c : E \rightarrow \mathbb{R}$, and $s, t \in V$, find an s–t-path whose longest edge is shortest possible.
Modify Dijkstra's algorithm in order to solve the bottleneck path problem. Prove the correctness of your algorithm. [10 + 8 = 18]

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination: 2012-2013

M. Tech. (CS) 2nd Year

Artificial Intelligence

Date: 19.09.2012 Maximum Marks: 60 Duration: 2 hours

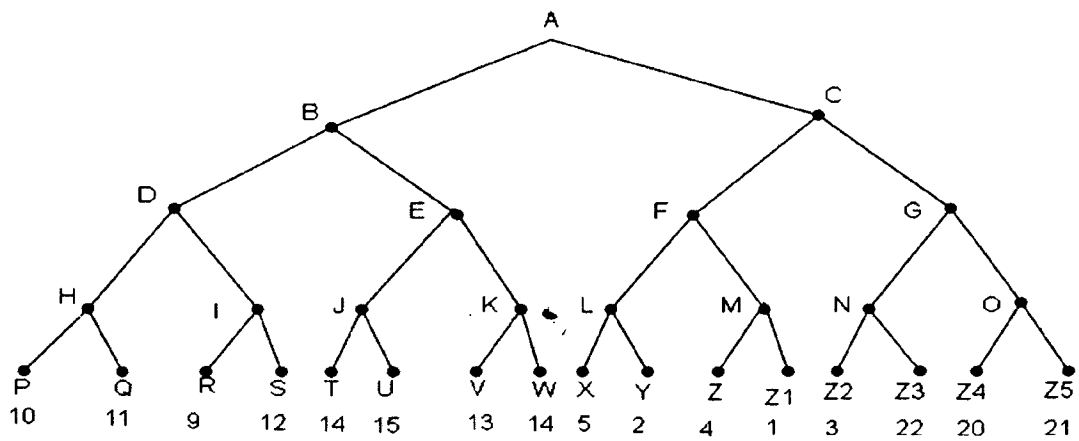
Answer all questions in brief.

1. In *farmer-fox-goose-grain* puzzle, a farmer wishes to cross a river taking his fox, goose, and grain with him. He can use a boat which will accommodate only the farmer and one possession. If the fox is left alone with the goose, the goose will be eaten. If the goose is left alone with the grain it will be eaten. Draw a state space search tree for this puzzle using left-bank and right-bank to denote left and right river banks, respectively. [10]

2. Prove that any monotonic heuristic is *admissible*. Describe the *crossover* and *mutation* operators of genetic algorithm. [4 + (3+3) = 10]

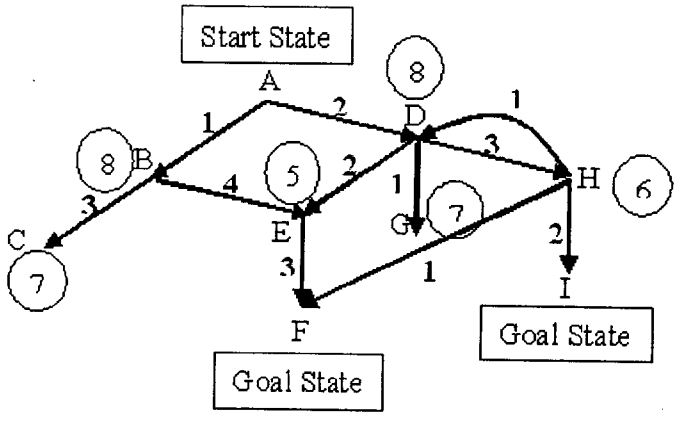
3. Describe the *depth-first iterative deepening* algorithm. Prove that it is asymptotically optimal among brute-force tree searches in terms of time, space, and length of solution. [4 + (3 + 2 + 1) = 10]

4. Perform the *minimax* search procedure on the game tree shown below in which the static scores are all from the first player's point of view and MAX is allowed to move first. Perform the left-to-right α-β pruning procedure on this tree and show how many nodes can be pruned away. [4 + 6 = 10]



5. Execute the *uniform cost search* and *best first search* algorithms on the following search graph, and show the solution path, along with its cost and list the expanded nodes for each

case (each node of the graph is represented by a letter and the encircled value is the heuristic evaluation of the corresponding node, while the bolded numerical value represents the actual length of the path between two nodes). [5 + 5 = 10]



6. Solve the following cryptarithmetic problem:

```
  S E N D
+ M O R E
-----------
M O N E Y
-----------
```

[10]

INDIAN STATISTICAL INSTITUTE

M. Tech (Computer Science) II year, 2012 – 13

Pattern Recognition and Image Processing

Date: 22.09.12 **Maximum marks: 60** **Duration: 150 minutes**

Note: Answer all the questions

1. State the Bayes decision rule for three-class classification problem and show that it minimizes the probability of misclassification. [3+10=13]
2. Let there be two classes C_1 and C_2 with prior probabilities *0.5* and *0.5*, and class conditional density functions p_1 and p_2 where

$$p_1(x) = 3e^{-3x};\ 0 < x < \infty,$$
$$= 0 \text{ otherwise,}$$

and $p_2(x) = 2e^{-2x};\ 0 < x < \infty$
$$= 0 \text{ otherwise.}$$

(i) Find the Bayes decision rule for the above classification problem and find its probability of misclassification.

(ii) Find the probability of misclassification for the following decision rule and verify that it is not less than the misclassification probability of the Bayes decision rule.

Put x in class 1 if $x \geq 2$. Otherwise put it in class 2.

[(5+5)+(5+5)=20]

3. (i) State the minimum within cluster distance criterion.
(ii) Describe the c-means algorithm for clustering.
(iii) Give an example of a data set and two of its initial partitions for which the resultant clusterings would be different when c-means algorithm is applied.

[5+5+7=17]

4. (i) State the minimum distance classifier.
(ii) Derive the minimum distance classifier from the Bayes decision rule.
(iii) State the k nearest neighbor rule for pattern classification. [2+5+3=10].

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination: 2012-2013
M Tech (CS) 2nd year
Internet and Multimedia Technologies

Date: 24.09.2012 Maximum Marks: 50 Duration: 90 minutes

Answer any TWO out of questions 1, 2 and 3 and answer any SIX out of questions 4, 5, 6, 7, 8, 9, 10 and 11.

1. Write TRUE or FALSE [10]:

a) Signal frequency information is required to draw the line spectra of the signal.
b) Energy of an audio signal is a function of the square of the amplitude of the signal.
c) High pass filters blocks high frequency information.
d) MPEG is an image compression scheme.
e) CMY to YCbCr transformation is required for printing images on white paper.
f) Cosine transform can transform time domain signal to frequency domain.
g) Sign of audio amplitude changes at zero-crossings of the audio signal.
h) Vanishing point is a 2D point on the optic axis.
i) Fundamental frequency of a signal is always a low frequency signal.
j) A 24bit color image can have maximum 65536 numbers of colors.

2. Given coordinates of two points (2,4) and (8,6) find out all the intermediate points lying on the digital line displayed in a computer screen connecting points (2,4) and (8,6). Follow scan conversion principle to draw the digital line between the points. [10]

3. A circle of diameter 20 pixels is drawn in the center of the computer screen having dimension 320x240 pixels. Give algorithmic steps to transform the circle into an ellipse having major and minor radii of 40 pixels and 15 pixels respectively. Note that the ellipse is to be displayed such that the center of the ellipse now lies in the screen coordinate (100, 100). [10]

4. We need to code a sequence f=11, 12, 17, 15, 12. Define a predictive coding scheme for the above sequence for which compression can be achieved. Show that the lossless decoding can be achieved for the proposed predictive coding scheme. (2+3=5)

5. (a) Use a dictionary based coding scheme to compress the input string ABBABBBABCABA. Discuss the extent of compression achieved for the proposed coding scheme.
(b) State two properties of Cosine Basis Functions that ensure better decomposition of cosine transformed signal. (3+2=5)

6. Assume N bits per sample are used for quantization. What is the maximum quantization error for a particular sampling time? For each additional bit available for quantization, what would be the additional signal to noise ratio value? Define non-linear quantization. (1+2+2=5)

7. Draw the block diagram of MPEG-Audio Layer III codec scheme and state the purpose of each block in one line. (5)

8. Specify four time domain features of an audio signal. How RGB color of a pixel value can be transformed to HSV space? (2+3=5)

9. Derive an expression based on which one single image mask (a sub-image) can be designed that can smooth an image and subsequently detect the high frequency information in the image. (5)

10. In a given imaging application we like to count the number of objects present in the gray level image. State implementation details of the solution to the problem. (5)

11. Derive 2D image coordinates for a 3D object as seen by a viewer. Assume that the coordinates for the 3D object, 2D image plane and the viewer are defined in the 3D world coordinate system. (5)

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination : (2012-2013)

M.Tech. (Computer Science) II Year.

Database Management Systems

Date: 29.09.2011 **Maximum Marks: 60** **Time: 2 Hours.**

1. Part of a university database wants to relate a set of Faculty members with Project-linked students along with the research projects conducted by those faculties. Consider the following information about the entities involved.
   a) Each faculty member has an employee no (unique), a name, a designation (Prof, Asso.Prof etc) and a research area.
   b) Each research project has a project no (unique), a sponsor name (DST, CSIR etc), a start date, an end date and a budget.
   c) Each project-linked student has a roll no (unique), a name and a degree program (M.Sc, M.Tech, Ph.D etc). Each such student will work in only one project.
   d) One or more faculty members are associated with each project.
   e) Each project is managed by one faculty member known as its principal investigator.
   f) A faculty member may manage and/or work in more than one project.
   g) Each faculty member and each student is associated with some department.
   h) Each department is identified by a unique name. A department also has an address and a budget.
   i) Each project is monitored by one of the departments.
   j) A student works only in such a project that is monitored by the department where he/she is associated with.
   k) A faculty member, however, may work in a project which is not monitored by his/her own department.

   From the above description, draw an appropriate ER/EER diagram.
   To design a relational schema, derive a set of relations from the ER/EER diagram using the standard mapping rules.
   (15+15=30)

2. Add to the problem given in Question 1 that each project-linked student who is working for Ph.D will have a faculty member associated, known as the supervisor and the concerned student and the corresponding supervisor belong to the same department. Indicate what changes to be done in the ER/EER diagram drawn earlier and the corresponding changes in the relations derived.
   (5+5=10)

3. a) Two relations R_1 (a,b,c,d,x) and R_2 (a,b,c,d,y) have four common attributes. The fifth attribute x and y of R_1 and R_2 respectively also belong to the same domain. Find the intersection of the two relations without using 'union', 'intersection' or 'set difference' operators.
   b) Let R_1 (A,B,C) and R (A,D,E) are two relations where A is the common attribute. Attributes B and E are in the same domain. Attributes C and D are also in the same domain. However, the domain of B and E is different from the domain of C and D. Out of the two equivalent algebraic expressions given below, which one you would prefer and why?
   i) $\prod_A (R_1 \cup R_2)$
   ii) $(\prod_A (R_1)) \cup (\prod_A (R_2))$ (7+3=10)

4. Person (p_name, p_st_name, age)
   Company (c_name, c_st_name, annual turnover)
   Worksfor (p_name, c_name, designation, salary)
   Managedby (p_name, m_name)

   For the relational schema shown above consider the following constraints.
   a) Name of a person (p_name) is unique.
   b) Company name (c_name) is also unique.
   c) Every person working in a company reports to a manager identified by his/her name (m_name).
   d) A manager being a person would also have an entry in the person relation.
   e) p_st_name and c_st_name signify street name,

   Using relational algebra form the query:
   Find the list of p_name where a person resides in the same street as his/her manager.
   (10)

INDIAN STATISTICAL INSTITUTE

First Semester Examination: 2012-13

Course Name: M. Tech. (CS) 2nd Year

Subject Name: Multimedia and Internet Technologies

Date: 30.11.2012 Maximum Marks: 100 Duration: 3 hours

Note: Answer all questions in both Part-I and Part-II
Use *separate* answer sheets for Part-I and Part-II (*do not mix up*)

Part-I: Multimedia (Marks: 35)

A1. (i) Given a 2D digital signal matrix

$$\begin{bmatrix} 3 & 5 & 7 \\ 2 & 7 & 6 \\ 3 & 4 & 9 \end{bmatrix}$$

and convolution mask

$$\begin{bmatrix} 1/4 & 1/2 & 1/4 \\ 1/2 & 1 & 1/2 \\ 1/4 & 1/2 & 1/4 \end{bmatrix}$$

what will be the digitized signal value at the centre of the matrix (whose value is 7 prior to convolution) after convolution?

[2]

(ii) Find the bit budget per minute for 16-bit (stereo) audio digitized at 44.1 kHz and compressed at 12:1 ratio.

[3]

A2. (i) How between-frame motion vectors are estimated and encoded for MPEG-I video?

[4+3]

(ii) What are the decompression steps to retrieve RGB values for a JPEG image?

[3]

A3. How a circle is drawn in a digitized pixel grid using Bresenham's algorithm?

[10]

A4. (i) Name three schemes for low-pass signal filtering.

[3]

(ii) How arithmetic code of an input symbol set is calculated? Give an example.

[7]

Part-II: Internet Technologies (Marks: 65)

B1. Match the following (10 marks):

(a) MapReduce	i. Java Servlet
(b) Spoofing	ii. Struts framework
(c) SVG (Scalable Vector Graphics)	iii. Affinity
(d) DTD	iv. Network Security
(e) On the fly webpage creation	v. XMPP
(f) Event driven asynchronous I/O	vi. Social network
(g) Blogging	vii. Cloud Computing
(h) Edgerank	viii. HTML 5
(i) Model-View-Controller	ix. XML document
(j) Instant messaging	x. Node.js

B2. Short answer (30 marks):

(i) What do you mean by Cookies? How this concept is useful in web programming?

[2+3 = 5]

(ii) Write down two significant differences between

(a) XML and HTML

(b) JSP and Java Script

[3+3 = 6]

(iii) Give two specific types of applications that can benefit from using XML

[4]

(iv) Why do I need JSP technology if I already have servlets?

[3]

(v) What do you understand by *structured*, *unstructured* and *semi-structured* data?

[3]

(vi) What type of difference will Facebook and Twitter induce if their contents are modelled as graphs?

[3]

(vii) Explain the screened host architecture for setting up a firewall.

[3]

(viii) List down two significant threats in the context of data security in cloud computing.

[3]

B3. Problems (25 marks):

(i) Write a XML schema that provides tags for a person's first name, last name, weight, and shoe size. Weight and shoe tags should have attributes to designate measuring systems.

[5]

(ii) Explain the functionalities of the following PHP programs/code segments. Wherever possible show the output.

(a)

```
<?php
//connection stuff goes here
$sql = "SELECT field FROM table";
$result = mysql_query($sql);
$arrContent = mysql_fetch_array($result);
echo "<a href=\"http://fsck.dk/index.php/?field=".
$arrContent["field"]."\">Click here</a>";
?>
```

(b)

```
<?php
$numbers = array("1.gif","2.gif","20.gif","10.gif");
natsort($numbers);
print_r($numbers);
?>
```

[3+3=6]

(iii) Design an algorithm to determine whether two persons are connected in a large social network (e.g. Facebook). Note that all the profiles are not stored in a single machine. Write a pseudo code for your algorithm and briefly explain the steps. Clearly state your assumptions.

[6]

(iv) How is the impact factor (IF) of a journal counted? [Say, you are asked to calculate IF of a journal in 2012 using previous two years" data.]

[3]

(v) Using the idea of forward and back links explain the simplified PageRank algorithm.

[5]

Indian Statistical Institute

Semester-1 2012-2013

M.Tech.(CS) - Second Year

End-semester Examination (26 November, 2012)

Subject: Compiler Construction

Total marks: 125 Maximum marks: 120 Duration 3.5 hrs.

Please keep your answers brief and to the point.

Consider the following syntax-directed definition (SDD):

$E \rightarrow E_1$ **or** T	{ E1.true = E.true; E1.false = newlabel(); T.true = E.true; T.false = E.false; }
$E \rightarrow T$	{ T.true = E.true; T.false = E.false; }
$T \rightarrow T_1$ **and** F	{ T1.true = newlabel(); T1.false = T.false; F.true = T.true; F.false = T.false; }
$T \rightarrow F$	{ F.true = T.true; F.false = T.false; }
$F \rightarrow$ **id**$_1$ **relop id**$_2$	{ gen(id1.name, id2.name, relop.op, F.true); gen(F.false); }

(a) Is the above SDD (i) S-attributed? (ii) L-attributed? Justify your answer in 1-2 lines.

(b) Convert the SDD into an equivalent translation scheme (TS) by embedding actions at appropriate places in the RHS of each rule.

(c) Convert the TS into a form suitable for bottom-up parsing by using markers where necessary. *Avoid introducing unnecessary markers.*

(d) Compute I_0, the set of items corresponding to the start state for an SLR parser for the grammar in part (c).

(e) Use (d) to show that the above SDD **cannot** be handled by a one-pass augmented SLR parser that executes semantic actions while parsing an input string. [3 + 6 + 8 + 6 + 5 = 28]

Consider the following SDD (similar to the one given in question 1).

$E \rightarrow E_1$ **or** $M\ T$	{ bp(E1.flist, M.place); E.tlist = E1.tlist + T.tlist; E.flist = T.flist; }
$E \rightarrow T$	{ E.tlist = T.tlist; E.flist = T.flist; }
$T \rightarrow T_1$ **and** $M\ F$	{ bp(T1.tlist, M.place); T.tlist = F.tlist; T.flist = T1.flist + F.flist; }
$T \rightarrow F$	{ T.tlist = F.tlist; T.flist = F.flist; }
$F \rightarrow$ **id**$_1$ **relop id**$_2$	{ F.tlist = 1; F.flist = 0; }
$M \rightarrow \varepsilon$	{ M.place = nextlocation(); }

(a) Eliminate left-recursion and left factors (if necessary) from the grammar in the above SDD.

(b) Construct an equivalent translation scheme (TS) that uses the modified grammar obtained in part (a). Make sure attributes are handled correctly. HINT: Draw the parse tree for a suitable example string using the original grammar. Annotate the nodes with appropriate attribute values. Then construct the parse tree for the same string using the non-left-recursive version of the grammar. Use this parse tree to determine what semantic actions are needed. If your answer is incorrect, you will get partial credit if you show these steps clearly.

[4 + 15 = 19]

P.T.O

3. (a) Write a grammar to represent the syntax of C structure definitions. Assume that the fields can only be of `char`, `int`, or `float` type, or can themselves be structures. You do not need to consider pointers, arrays, etc. Your grammar should permit the following structure definition.

```
struct foo {   char c;   struct point { float x, y; } p;   }
```

(b) Add semantic rules to your grammar to calculate the byte offset for each field in a structure definition. You may ignore alignment restrictions. Assume that (i) the fields of a structure are stored in contiguous locations, and in the order in which they are declared; (ii) the sizes of `char`, `int` and `float` fields are 1, 4 and 4 bytes respectively; (iii) there is a suitable auxiliary function (say, `store_offset()`) that stores the calculated byte offsets in the appropriate location in the symbol table.

(c) Write a grammar for assignment statements that permits the use of structures in addition to simple variables. Again, you may ignore arrays, pointers, etc.

(d) Using your grammar, draw the parse tree for: `a.p.x = 10.0`

[8 + 13 + 8 + 4 = 33]

4. Consider the following function in C.

```
int subseq_search(int *seq, int L, int *sub, int l)
{   int i, j;
    for (i = 0; i < L - l + 1; i++) {
        for (j = 0; j < l; j++) {
            if (seq[i+j] != sub[j])
                break;
        }
        if (j == l) return i;
    }
    return -1;
}
```

(a) Convert the body of the procedure into 3-address code. Each time a temporary variable is needed, use a new temporary. You may use the name of an array instead of the constant (base address) associated with that array. **Do not perform any optimization at this stage.**

(b) Write the machine code for the calling sequence and return sequence for a call to `subseq_search` from `main`. Assume that (i) the stack grows from low addresses to high addresses; (ii) the stack pointer points to the beginning (i.e. lowest address) of an AR; and (iii) the AR for `main` occupies 320 bytes.

(c) Identify the leaders (and thus the basic blocks) in the 3-address code in (a).

(d) Optimize your intermediate code in (a) by using whichever of the following techniques are applicable: constant folding, global common sub-expression elimination, copy propagation, dead code elimination, code motion, induction variable elimination.

[13 + 10 + 8 + 14 = 45]

END

INDIAN STATISTICAL INSTITUTE

Semester Examination : (2012-2013)

M.Tech.(CS) II Year

Database Management Systems

Date:28.11.2012 **Maximum Marks: 50** **Duration: 2 Hours**

Answer all questions

1. For the relation R=(A,B,C,D,E,F), the following dependencies are specified :

   A →BC, BC →D, CD →E, EA →F

   Considering the above dependencies,

   a) Derive all the possible candidate keys of R.
   b) Explain whether the relation R is free from partial and transitive dependencies.
   c) If necessary, decompose R into a set of normalized relations removing the partial and transitive dependencies.
   d) If a new multivalued dependency D->->F is introduced replacing the functional dependency EA →F, what would be the new set of normalized relations?

   (4x4=16)

2. Three transactions T_0, T_1 and T_2 are executed in the way as shown below. All the three transactions are manipulating the same data item A. If crash occurs in one of the six places (1 to 6) as indicated in the schedule, explain the recovery action the system would undertake if it follows a deferred update log maintenance strategy with standard 'redo' and 'undo' routines when the log contains no check point. Now, if a check point is inserted in the log after the commit of T_0, would there be any change in the recovery process after the crash points 5 and 6.

   Schedule:

   T_0 : read (A)
   A=A-1
   write(A)
   ----------(1)
   T_1 : read (A)
   A=A+2
   write(A)
   ----------(2)
   commit
   ----------(3)
   T_0 : commit
   ----------(4)
   T_2 : read (A)
   A=A*5
   write(A)
   ----------(5)
   commit
   ----------(6)

   (8x2=16)

3. Two relations **R** and **S** are to be joined against a common attribute **a** where **a** is the primary key of **S**. **R** has 10000 tuples and S has 2000 tuples.

   a) If for both the relations 100 tuples form a page, what would be the estimated number of disk accesses including both reading of relations and writing of results. Consider that the main memory can accommodate 30 blocks and block-oriented nested loop join is used as the join algorithm.
   b) If each tuple in relation **R** and **S** are of length 100 bytes and 200 bytes respectively and the common attribute **a** is 30 bytes long, find the estimated size of the joined relation in bytes.
   c) If the common attribute **a** is not a key of either **R** or **S**, what would be the estimated size of joined relation in bytes when **V(a, R)**=200 and **V(a, S)**=100 and all other data are same as earlier.

   (5+6+7=18)

INDIAN STATISTICAL INSTITUTE
M. Tech.(Computer Science) II Year, 2012-13
Semestral Examination
Pattern Recognition and Image Processing

Date: 03·12·12 Maximum Marks: 100 Duration: 195 minutes

Note: This paper carries 111 marks. Answer as much as you can.

1. Apply the histogram equalization method for the following frequency distribution of gray values. [6]

Gray value	Frequency
0	1
1	10
2	5
3	27
4	17
5	2
6	16
7	2

2. Describe a method for introducing salt and pepper noise randomly in a gray level image. [5]
3. Describe the Canny edge detection algorithm for gray level images. [10]
4. Describe the Hough Transform method for detecting line segments in a binary image. [12]
5. Describe, along with the underlying principles, a region based segmentation technique for images [12]
6. Describe a thinning method for a 4-connected binary component in an image. [10]
7. Using the directional numbers for 8-directions, describe the boundary of the object shown below. The given binary image has 8 rows and 7 columns, and "*" denotes object pixel. [6]

		*	*	*	*	
	*	*	*	*	*	*
		*	*	*	*	
	*	*	*	*	*	
*	*	*	*	*	*	
*	*	*	*	*	*	*
	*		*	*	*	*
					*	

8. Draw the solution tree for branch and bound feature selection algorithm if 3 features are to be selected from 7 features. [6]

(P.T.O)

9. Let $\underline{X}' = (X_1, X_2, X_3)$ be a random vector with dispersion matrix $\begin{pmatrix} 2 & 0 & -1 \\ 0 & 8 & 0 \\ -1 & 0 & 3 \end{pmatrix}$.

Find the two principal components of $\underline{X}$. [8]

10. Describe any four feature selection algorithms. [4x4=16]
11. Write short notes on the following. [4x5=20]
    (a) Probabilistic separability measures for feature selection.
    (b) Interclass distance based criterion functions for feature selection.
    © Median filtering of images
    (d) Skeleton of an object in an image

INDIAN STATISTICAL INSTITUTE

First-Semester Examination: 2012-2013

M. Tech. (CS) 2nd Year

Artificial Intelligence

Date: 05-12-12 Maximum Marks: 100 Duration: 3 hours

Answer all questions in brief.

1. Answer the following:
   a) Write a program in Prolog for merging two ordered lists.
   b) Explain with example the difference between red cut and green cut in Prolog.
   c) Describe the difference between the following two codes written in Prolog when the goal query is "grandfather(james, X)":

grandfather(X,Y):- father(Z, Y), father(X, Z). father(james, robert). father(mike, william). father(william, james). father(robert, hency).	grandfather(X,Y):-father(X, Z), father(Z, Y). father(james, robert). father(mike, william). father(william, james). father(robert, hency).

   d) Write a program in Prolog for post-order traversal of a binary tree. The traversal method stores the elements of the tree in a list. [5 + 4 + 5 + 6 = 20]

2. Answer the following:
   a) Prove that any monotonic heuristic is admissible.
   b) Describe the simulated annealing approach.
   c) Consider the following set of sentences. "Mary will get her degree only if she registers as a student and passes her examination. She has registered herself as a student. She has passed her examination." Prove using both i) semantic tableaux approach and ii) resolution refutation method that "she will get a degree". [4 + 6 + (5 + 5) = 20]

3. Answer the following:
   a) Prove that a clause C is a logical consequence of a set of clauses S if and only if the set $S' = S \cup \{\sim C\}$ is unsatisfiable.
   b) Prove using semantic tableaux approach that the following sentences are mutually consistent. "All Indian citizens who are adult have right to vote in election. Mary is an Indian citizen and has voting right. Mary is an adult."
   c) What are the main features of an expert system? Explain with an example the backward chaining approach in a rule based expert system. [6 + 6 + (4 + 4) = 20]

4. Answer *any two* from the following: [2 X 10 = 20]
   a) Discuss the Bayes' theorem for the probabilistic reasoning. Suppose an initial observation S_1 confirms some hypothesis h with the belief MB = 0.3. The second observation S_2 performs the same hypothesis h with the belief MB = 0.5. Find the certainty factor regarding the hypothesis h by the two observations S_1 and S_2. [6 + 4 = 10]
   b) Describe the following with suitable examples:
      (i) Dempster-Shafer theory of evidence;
      (ii) Means-ends analysis. [5 + 5 = 10]

c) Explain, with suitable examples, the differences between
(i) Knowledge-based system and expert system;
(ii) Uniform cost search and best first search. [5 + 5 = 10]

5. Answer the following:

a) Let $I = <U, A>$ be a decision table, where $U = \{x1,....., x7\}$ is a nonempty set of finite objects, the universe, and $A = C \cup D$ is a nonempty finite set of attributes. Here, C = {A1, A2} and D are the set of condition and decision attributes, respectively.

U	A1	A2	*D*
x1	16-30	50	yes
x2	16-30	0	no
x3	31-45	1-25	no
x4	31-45	1-25	yes
x5	46-60	26-49	no
x6	16-30	26-49	yes
x7	46-60	26-49	no

In the context of rough set theory, explain lower approximation, upper approximation, boundary region, and degree of dependency and significance of an attribute with the above example data. [2 X 5 = 10]

b) Consider a sliding block puzzle with the following initial configuration:

W	W	W	B	B	B	E

There are three white tiles (W), three black tiles (B), and an empty cell (E). The puzzle has the following moves:

(i) A tile may move into an adjacent empty cell with unit cost.

(ii) A tile may hop over at most two other tiles into an empty cell with a cost equal to the number of tiles hopped over.

The goal of the puzzle is to have all the black tiles to the left of all the white tiles without regard for the position of the empty cell. Define the problem as a state space graph problem and find a sequence of moves that will transform the initial configuration to a goal configuration. What is the cost of the solution? [4 + 6 = 10]

INDIAN STATISTICAL INSTITUTE

Semestral Examination

M.Tech(CS)-II Year, 2012-2013 (Semester-III)

Advanced Algorithms for Graphs and Combinatorial Optimization Problems

Date: 5.12.12 Maximum Marks: 100 Duration: **3.5 Hours**

Note: The question paper is of 117 marks. You can answer as much as you can, but the maximum you can score is 100.

---

1. Let $G = (V, E)$ be a directed acyclic graph and let k be the maximum number of edges in a path of G. Design an algorithm to divide the vertices into at most $k+1$ groups such that for every pair of vertices v and w in the same group there is no path from v to w and there is no path from w to v. The algorithm should run in linear time.
(**Hint:** *You may use topological sorting*) (15)

2. Design a backtracking algorithm for the graph isomorphism problem. Comment on the complexity of the algorithm. (10+2=12)

3. Let $G = (V, E)$ be a directed graph. Each edge (u, v) in E has an associated value $f(u, v)$, which is a real number in the range $[0, 1]$ that represents the reliablility of a communication channel from vertex u to vertex v. We interpret $f(u, v)$ as the probability that the channel from u to v will not fail and we assume that these probabilies are independent. Give an efficient algorithm to find the most reliable path between two given vertices. (20)

4. Design and analyze a linear-time algorithm for finding a maximum matching in a tree. (10)

5. Describe Edmonds' Blossom shrinking algorithm. Prove that when the algorithm terminates, it outputs a maximum matching for the current graph. (10+10=20)

6. Let G be a biconnected graph with an edge (s, t). Prove that G has an s-t order $v_1, \ldots, v_n$ with $v_1 = s$ and $v_n = t$. (10)

7. Let $G = (V, E)$ be an interval graph such that the generating intervals on the real line are known. Design an efficient algorithm to find a maximum independent set in G. (12)

8. Design and analyze a dynamic programming algorithm for the *traveling salesman problem (TSP)* for undirected weighted graphs. (12)

9. Prove that the vertex-cover problem remains *NP-complete* even if all the vertices in the graph are restricted to have even degree. (6)

INDIAN STATISTICAL INSTITUTE

Final Examination:(2012-2013)

MTech C.S. 2nd Year

Digital Signal Processing

Date: 07.12.2012 Maximum Marks: 100 Duration: 3 hours

Note: The marks add up to 115. The maximum you can score is 100. The exam is open-book, open-notes. Use of calculators is permitted.

1. Let $x[n]$, $0 \leq n \leq 4$ be a sequence of length 5 have a 10-point DFT $X[k]$.

   (a) Determine the DFT of the sequence $v[n] = \begin{cases} 0 & 0 \leq n \leq 4, \\ x[n-5] & 5 \leq n \leq 9 \end{cases}$

   (b) Consider a sequence $y[n] = x[n] + x[n-5]$. Determine $Y[k]$, the 10-point DFT of $y[n]$ in terms of $X[k]$.

   [5+5]

2. Consider a sequence $x[n] = \{1\ 2\ 3\ 4\ 5\ 6\}$, $0 \leq n \leq 5$. Let $X[k]$ denote the 6 point DFT of $x[n]$. Without actually computing the DFT or IDFT, determine

   (a) $y_1[n]$, the IDFT of $Y_1[k] = X(e^{j\omega})|_{\omega=2\pi k/4}$, $k = 0, \cdots, 3$.

   (b) $y_2[n]$, the IDFT of $Y_2[k] = W_3^{2k} X[k]$.

   [6+6]

3. A causal LTI system produces an output $y[n] = (0.5)^n u[n] - 0.4(0.2)^{n-1} u[n-1]$ for an input $x[n] = (0.6)^n u[n]$.

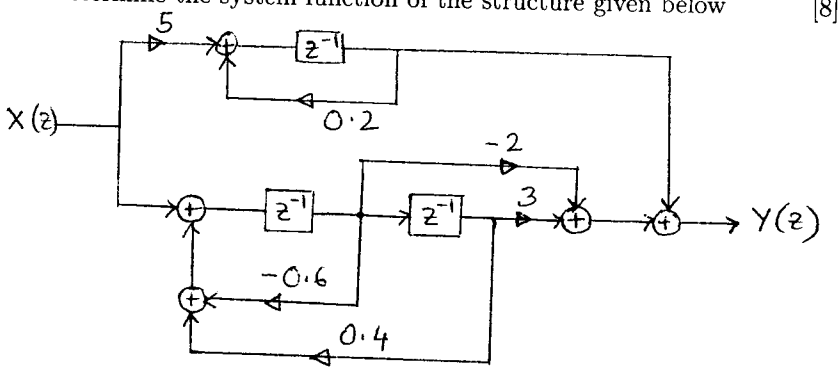
   (a) Obtain the corresponding difference equation.

   (b) Determine the impulse response of the system.

   (c) Draw a canonic Direct-form and a parallel realization of the system

   [5+5+(5+5))

4. Determine the system function of the structure given below [8]



5. Obtain the Discrete-time Fourier Transform (DTFT) of the sequence

$$x[n] = \begin{cases} \alpha^{|n|}, & |n| \leq 5, \\ 0, & \text{otherwise} \end{cases}$$

[5]

6. Determine the inverse DTFT of

$$X(e^{j\omega}) = \frac{e^{j\omega}(1 - e^{j\omega N})}{1 - e^{j\omega}}$$

where N is an integer. [5]

7. Determine if the following are stable systems, giving proper justification

   (a) A cascade connection of two stable LTI systems.

   (b) A parallel connection of two stable LTI systems.

[5+5]

8. The system function of a discrete-time LTI system is given by:

$$H(z) = \frac{3(z + 1.1)}{(z - 0.7)(z + 2)}$$

   (a) Can this system be both stable and causal?

   (b) Determine the impulse response of the system. Does the frequency response of this system exist? [5+(5+2)]

9. A Type-I filter with real coefficients has zeros at 1, -0.6 and -1+j. Determine the locations of the remaining zeros and the corresponding transfer function. [5+3]

10. Determine the z-transform of $x[n] = n^2 0.6^n u[n]$. Indicate the ROC (region of convergence). [5]

11. Determine the inverse z-transform of

$$X(z) = log(\frac{\alpha - z^{-1}}{\alpha}),\ |z| > 1/|\alpha|,\ \alpha < 1$$

[8]

12. The linear convolution of two real sequences of lengths 16 and 7, respectively is to be performed. Determine the least number of real multiplications needed for each of these methods:

    (a) Direct computation of the linear convolution.

    (b) Using a single circular convolution.

    (c) Using a radix-2 FFT algorithm. Do not include multiplications by ± 1, $\pm j$, and W_N^0

[3+3+6]

Indian Statistical Institute

M. TECH. (CS) 2 Year : 2012–2013
Semester Examination
Subject: Cryptology

Date: 07/12/2012 Time: 3 hours Maximum Marks:100

Note: The paper carries 110 marks. Maximum you can score is 100.
Notations used are as explained in the class.

1. Give the definition of random oracle model. Prove that under some reasonable assumptions, a collision resistant hash function is preimage and second preimage resistant. [4 + 11 = 15]

2. Give an algorithm for the Merkle-Damgard construction. Prove that the construction gives a collision resistant hash function when the compression function used in it is collision resistant. [5 + 10 = 15]

3. Define the RSA cryptosystem. Prove that in RSA the following holds $d(e(x)) = x$, for any $x \in \mathbb{Z}_n$. Prove that RSA is not IND-CPA secure. Prove that ElGamal is not IND-CCA secure. [3+7+2+3 = 15]

4. Describe Shank's algorithm to find discrete log. Give an example of a group where the discrete log problem is easy. Give an example of a group where the discrete log problem is believed to be hard but the decision Diffie-Hellman problem is easy (Justify). [8 + 2 + 5 = 15]

5. Describe the ElGamal signature scheme. Describe the various attack models and the goal of the adversary on a signature scheme. Prove that there is an existential forgery under key only attack on the ElGamal signature scheme. [5 + 5 + 5 = 15]

6. Explain Shamir's (t, w)-threshold scheme. Prove that it is a perfect secret sharing scheme. [5 + 10 = 15]

7. Let E be an elliptic curve $y^2 = x^3 + x + 6$ over $\mathbb{Z}_{11}$.
   (a) Determine the points on $E(\mathbb{Z}_{11})$.
   (b) Compute $(2, 7) + (2, 7)$ where $+$ is the addition operation defined on the elliptic curve.
   (c) Compute $(5, 2) + (2, 7)$.
   (d) Compute $-(2, 7)$.
   (e) Compute $(2, 7) + \mathcal{O}$ (where $\mathcal{O}$ is the point at infinity). [8 + 5 + 5 + 1 + 1 = 20]

INDIAN STATISTICAL INSTITUTE
End-Semestral Examination: 2012 – 13

Course Name: M. TECH CS - II

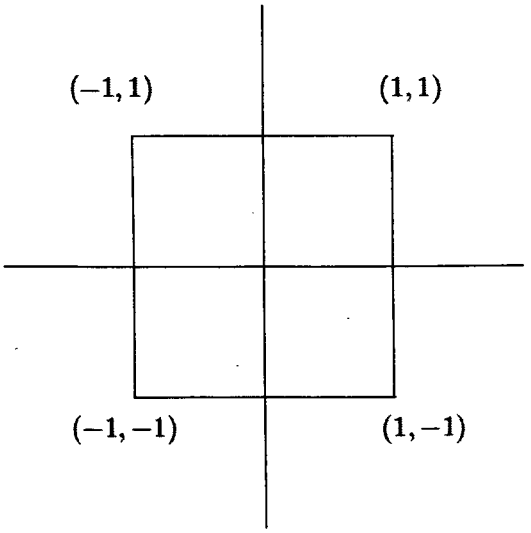
Subject: Computer Graphics

Date: 07.12.2012

Maximum Marks: 100 Duration: 3 hrs.

Answer as much as you can.

1. a) Here is a test pattern in the plane:



Draw its image after being transformed by each of the following (individually, not cumulatively):

(i)

$$\begin{bmatrix} \cos 30 & \sin 30 & 0 \\ -\sin 30 & \cos 30 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(ii)

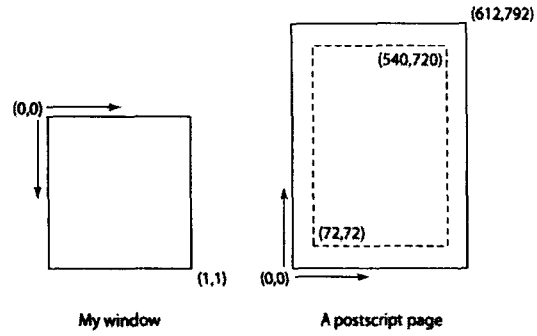
$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(iii)

$$\begin{bmatrix} 1 & 0 & 0.2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

(4+3+3 = 10)

b) I have a picture that I've expressed in a normalized 2D coordinate system with top left corner $(0,0)$ and bottom right corner $(1,1)$ (y grows downward, as in screen coordinates). I'd like to print the picture as PostScript, centered on an 8.5" × 11" page with a one-inch margin. One unit in PostScript is equal to $\frac{1}{72}$". The picture should be centered on the page, and scaled uniformly to be as large as possible while still fitting within the margins. The two coordinate systems are shown below.



Write a short C function or give a matrix that takes normalized coordinates and outputs PostScript coordinates.

(10)

2. a) Given a clipping window P(0, 0), Q(30, 0) R(30, 20), S(0, 20), use Sutherland-Cohen algorithm to determine the visible portion of the line A(10, 30) and B(40, 0). (8)

b) Define the frustum of vision to be centered on the Z-axis with the eyepoint at (0, 0, -*d*) and with the window on the x, y – plane having its center at (w_{cx}, w_{cy}, 0), width $2*w_{sx}$, and height $2*w_{sy}$. The far clipping plane is to be parallel to the window and intersect the Z-axis at (0, 0, *f*).

i) Give the equations (in the window coordinate system) of the six clipping planes. Adjust the signs in your equations (if necessary) so that a positive value denotes "visible" and a negative value denotes "invisible".

ii) What is the perspective matrix appropriate for this formulation of the problem? Explain your answer.

(7+5)

3. a) Two cubic Bezier curves, with control points P_0, P_1, P_2, P_3 and Q_0, Q_1, Q_2, Q_3 respectively, are joined together at $P_3 = Q_0$. What must be true for them to meet with C^1 continuity?

(2)

b) A cubic Bézier curve has the mathematical form

$$C(t) = \sum_{i=0}^{3} P_i b_i(t)$$

i) Give explicit formulas for $b_0(t), \ldots, b_3(t)$.

ii) Show that $b_i(t) \geq 0$ for $0 \leq t \leq 1$.

iii) Show that $\sum_{i=0}^{3} b_i(t) = 1$ for all t.

iv) Suppose $\mathbf{R}$ is a rotation matrix. Show that $\mathbf{R}C(t)$ is given by the Bézier curve that has control points $\mathbf{R}P_i$.

(12)

c) You are standing on the middle of the railroad tracks when you see that a train is approaching in the distance. You hold your thumb out in front of one eye at arm's length, and notice with interest that your thumb exactly covers the width of the train. You know that your arm is one metre long and your thumb is two centimetres wide. Looking down, you judge that the railroad tracks are 1.5 metres apart, and you see that the train's width extends exactly to the width of the tracks.

1. How far away is the train?
2. If the train is travelling at 50km/h, how many seconds do you have to get out of the way?

(4+2)

4. a) Discuss the diffused and specular reflections with respect to the basic illumination models.

(5)

b) Given a point for shading on a surface, we usually consider a normal vector $\vec{n}$, a vector to the light source $\vec{l}$, a vector to the viewer $\vec{v}$, and a reflection direction $\vec{r}$ (all vectors are of unit length). One possible version of the illumination equation is

$$I_{out}(\vec{v}) = k_a I_a + \sum_i \left(k_d + k_s \frac{P(\vec{r}, \vec{v})}{\vec{n} \cdot \vec{l}} \right) I_0^i \frac{\vec{l} \cdot \vec{n}}{c_1 + c_2 d_i + c_3 d_i^2}$$

i) Give an equation for the reflection direction vector $\vec{r}$ in terms of vectors $\vec{l}$ and $\vec{n}$.

ii) What is the term $k_a I_a$? How does increasing I_a affect the rendering of objects in the scene?

iii) What is the term $c_1 + c_2 d_i + c_3 d_i^2$? How is d_i calculated?

iv) This equation doesn't mention R, G, or B, and yet we use it to render coloured objects. How do we adapt the equation to handle RGB?

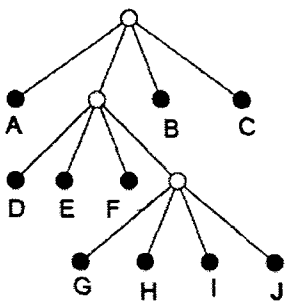
(4×3=12)

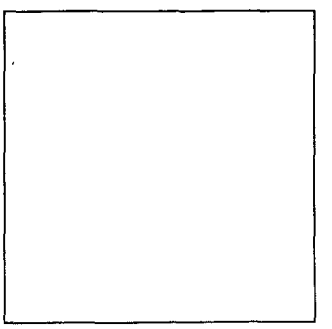
c) A ray $(P, \vec{v})$ hits a surface S at a point Q, where the normal to S at Q is $\hat{n}$. Draw a picture and give the formula for computing the reflected ray $\vec{r}$ emanating from Q.

(3)

5. a) The tree below represents a quadtree subdivision of a square. The leftmost branch is the upper left quarter, the next branch is the upper right, the third branch is the lower left, and the rightmost branch is the bottom left of a region. The leaf nodes are labelled A – J.

Subdivide the square below to indicate the subdivision of the quadtree. Label each portion of the square with the appropriate label from the tree.





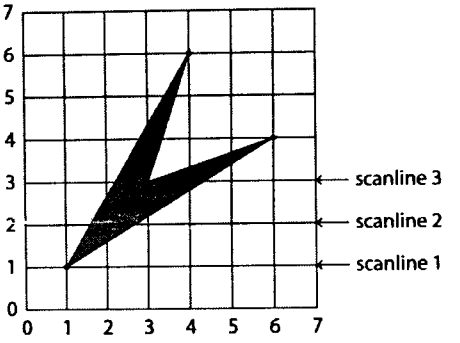
(5)

b) Explain the role of fractals in 3-D object representation. What do you mean by fractal dimension? How is it calculated? 10+2+3 = 15

6. a) Implement any one depth buffer method to display the visible surfaces of a given polyhedron. How can the storage requirements for the depth buffer be determined from the definition of the objects to be displayed? (10)

b) Give an algorithm for drawing the part of a circle which lies in the first octant. Assume that the circle has integer radius and is centered at the origin. Assume that you have a function *setpixel* (x, y) which turns on pixel (x, y). (10)

7. a) In the following, we want to apply the scan line algorithm for rasterization of random polygons in 2D to this example:



i) Write down the complete edge table (ET) that is used by this algorithm.
ii) Write down all the entries in the active edge table (AET) when the current scanline is at position 1, 2, and 3, respectively (cf. image). (5+5)

b) Outline a scheme for tweening by using spline interpolation. (10)

8. a) Derive expressions for converting RGB colour parameters to HSV values. (5)

b) How many different colour combinations can be generated using halftone approximations on a two-level RGB system with a 3 by 3 pixel grid? (3)

c) What do you mean by morphing? Explain the steps involved with the warping part of a morphing algorithm with a suitable example. (3+9)

INDIAN STATISTICAL INSTITUTE

Supplementary Semestral Examination

M.Tech(CS)-II Year, 2012-2013 (Semester-III)

Advanced Algorithms for Graphs and Combinatorial Optimization Problems

Date: Maximum Marks: 100 Duration: **3 Hours**

Note: The question paper is of 120 marks. You can answer as much as you can, but the maximum you can score is 100.

---

1. We are given a weighted, directed graph $G = (V, E)$ in which edges that leave the source vertex s may have negative weights, all other edge weights are non- negative and there are no negative-weight cycles. Argue that Dijkstra's algorithm correctly finds shortes paths from s in this graph. (15)

2. Design a backtracking algorithm for the graph isomorphism problem. Comment on the complexity of the algorithm. (13+2=15)

3. Describe and analyze the Ford-Fulkerson algorithm for solving the maximum-flow problem in a flow network. Describe how to solve the maximum-bipartite-matching problem using Ford-Fulkerson algorithm for solving the maximum-flow problem in a flow network. (15+5=20)

4. **(a)** Give an example of a weighted connected undirected graph $G = (V, E)$ and a vertex v such that the minimum-cost spanning tree of G is the same as the shortest-path tree rooted at v. Justify your answer.
   (b) Give an example of a weighted connected undirected graph $G = (V, E)$ and a vertex v such that the minimum-cost spanning tree of G is very different from the shortest- paths tree rooted at v. Can the two trees be compeltely disjoint ? Justify your answer. (5+5=10)

5. Describe Edmonds' Blossom shrinking algorithm. Prove that when the algorithm terminates, it outputs a maximum matching for the current graph. (15+5=20)

6. Let G be a biconnected graph with an edge (s, t). Prove that G has an s-t order $v_1, \ldots, v_n$ with $v_1 = s$ and $v_n = t$. (10)

7. We are given a weighted directed graph that has a negative weight cycle reachable from a source vertaex s. Give an efficient algorithm to list the vertices of one such cycle. (10)

8. Prove that *traveling salesman problem (TSP)* for undirected weighted graphs is NP-Complete. Design and analyze a dynamic programming algorithm for the *traveling salesman problem (TSP)* for undirected weighted graphs. (8+12=20)

INDIAN STATISTICAL INSTITUTE

SUPPLEMENTARY EXAMINATION 2012-13

Course Name: M. TECH CS - II

Subject: Computer Graphics

Date:

Maximum Marks: 100 Duration: 3 hrs.

Answer as much as you can.

1. a) Calculate a 4×4 matrix to rotate points about a vector A = [1 -2 1] by 60 degrees. Show your steps.

   b) In the following, assume Δ is a 2D triangle with vertices at $A = [1, 1]^T$, $B = [3, 2]^T$, and $C = [2, 3]^T$. Suppose $P = \alpha A + \beta B + \gamma C$ with $\alpha = 1/6$, $\beta = 1/3$, and $\gamma = 1/2$.

      1. Is $P = \alpha A + \beta B + \gamma C$ a valid affine combination? Justify your answer.
      2. Is the point P in the interior of the triangle? Justify your answer.
      3. Suppose points A, B, and C are transformed by an affine transformation T. Give an expression with numerical values for the weights that expresses $T(P)$ as an affine combination of $T(A)$, $T(B)$, and $T(C)$.
      4. Compute the coordinates of P in the same frame as A, B, and C.
      5. Compute the coordinates of P relative to the frame formed by the vector $B - A$ and $C - A$ with origin A.

   (5+15 = 20)

2. a) Write a simple function for scan conversion of lines with negative slopes. Explain your steps.

   b) Develop a scan conversion algorithm to draw an ellipse whose major axis is oriented at an angle specified by the user.

   (10+10=20)

3. a) Suppose an RGB raster system is to be designed using an 10 inch by 12 inch screen with a resolution of 150 pixels per inch in each direction. If we want to store 6 bits per pixel in the frame buffer, how much storage (in bytes) do we need for system the frame buffer?

   b) Determine the form of the transformation matrix for a reflection about an arbitrary line given by $y = mx + c$.

   c) Prove that a uniform scaling (s_x=s_y) and rotation form a commutative pair of operations.

   (6+7+7=20)

4. Write short notes on any 2: (10+10 = 20)

   a) Constructive solid geometry modelling

b) Morphing

c) Octree structures

d) Phong specular reflection model.

5. a) Consider the line which starts at (4, 13) and ends at (22, 3).

i) How many pixels are there on this line?

ii) With the DDA (Digital Differential Analyzer) algorithm, what will be the amount added to the secondary component each time through the loop (incrementing value)? Justify your answer.

b) Develop and implement an algorithm for clipping a line against a triangular window.

(10+10 = 20)

6. a) Write a pseudocode for the Z-buffer algorithms for determining visible surfaces of piercing polygons and polygons with holes.

b) (i) Estimate the 2D affine transformation matrix, $\mathbf{T} = \begin{bmatrix} \boldsymbol{F} & \boldsymbol{v} \\ 0^T & 1 \end{bmatrix} \in \mathbb{R}^{3\times3}$, given its action on three homogeneous points:

$$\begin{pmatrix}1\\1\\1\end{pmatrix} \xrightarrow{T} \begin{pmatrix}2\\1\\1\end{pmatrix}, \quad \begin{pmatrix}1\\-1\\1\end{pmatrix} \xrightarrow{T} \begin{pmatrix}2\\1\\1\end{pmatrix}, \quad \begin{pmatrix}-1\\-1\\1\end{pmatrix} \xrightarrow{T} \begin{pmatrix}0\\1\\1\end{pmatrix}.$$

(ii) What kind of transformation does this matrix represent?

(10+10 = 20)

7. a) Give a parametric representation $f(u, v)$ of a sphere centered at $(1, 2, 3)$ that has the following properties:

- $f(0, 0) = (2, 2, 3)$
- $f(1, 1) = (0, 2, 3)$
- $f(0.5, 0) = (1, 2, 4)$

and is oriented so that the surface faces outward. As usual, u and v both range from 0 to 1.

b) Derive the homogeneous matrix for a perspective projection with center of projection at (0,0,0) and projection plane x+2y+3z+4=0?

c) Explain the algorithm for scan-converting convex polygons. It is not necessary to write the code, a precise narrative description is sufficient.

(6+7+7 = 20)

8. a) Given is a plane $\begin{pmatrix} 1 \\ -2 \\ 2 \end{pmatrix} \bullet \mathbf{p} = 3$ and a triangle with the vertices $\mathbf{p}_0 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$, $\mathbf{p}_1 = \begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix}$ and $\mathbf{p}_2 = \begin{pmatrix} 3 \\ 1 \\ -1 \end{pmatrix}$. Does the triangle intersect the plane? Prove your answer.

b) Given is a polygon with the vertices (0,0), (0,1), (1,0) and (1,1). Draw in the answer space the shape obtained after transforming this unit square with the homogeneous matrix:

$$\mathbf{M} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos(90) & -\sin(90) & 0 \\ \sin(90) & \cos(90) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Explain your answer.

c) What are the window to viewport mappings for W = (0, 150, 50, 250) and V = (10, 50, 0, 80)?

(7+8+5 =20)

Indian Statistical Institute

Semester-1 2012-2013

M.Tech.(CS) - Second Year

Backpaper Examination (2012)

Subject: Compiler Construction

Total marks: 110 Maximum marks: 100 Duration 3 hrs.

Please keep your answers brief and to the point.

1. Consider the following syntax-directed definition (SDD) based on a left recursive grammar:

$L \rightarrow$ **id** [*Elist*]	{ `Elist.array = lookup(id.name);`
	`L.place = Elist.array;`
	`L.offset = Elist.place * elt_size(Elist.array);` }
$L \rightarrow$ **id**	{ `L.place = lookup(id.name); L.offset = 0;` }
$Elist \rightarrow Elist_1$, E	{ `Elist1.array = Elist.array;`
	`Elist.ndim = Elist1.ndim + 1;`
	`Elist.place = E.place +`
	`Elist1.place * num_elts(Elist1.array, Elist1.ndim);` }
$Elist \rightarrow E$	{ `Elist.ndim = 1; Elist.place = E.place;` }

(a) Explain in one line each whether `Elist.array`, `Elist.ndim`, `Elist.place`, `L.place`, `L.offset`, are inherited or synthesized attributes. Is the above SDD (i) S-attributed? (ii) L-attributed? Justify your answer in 1-2 lines.

(b) Convert the SDD into an equivalent translation scheme (TS) by embedding actions at appropriate places in the RHS of each rule.

(c) Convert the TS into a form suitable for bottom-up parsing by using markers where necessary.

(d) Assume that the attributes used in your TS are stored in the usual positions in the value stack during bottom-up parsing. For each semantic action in your TS, write code that manipulates the value stack according to the semantic action.

(e) Eliminate left-recursion and left factors from the grammar in the above SDD.

(f) Construct an equivalent translation scheme (TS) that uses the modified grammar (part (e)). Make sure attributes are handled correctly.

HINT: Draw the parse tree for the string `x[i,j]` using the original grammar, and look at the values of attributes at various nodes. Now construct the parse tree for the same string using the non-left-recursive version of the grammar.

[7+2+3+8+4+16=40]

2. In Pascal, the **repeat-until** statement has the form: `repeat S until E;`. It is equivalent to

```
do S while (not E);
```

(a) Manually translate the statement `repeat i = i+1 until (i == N);` into 3-address code.

(b) Give a syntax directed definition (SDD) for translating a general repeat-until statement into 3-address code using (i) symbolic labels, (ii) backpatching. You do not need to include grammar rules / semantic actions for general statements / Boolean expressions.

[2+(4+6)=12]

P.T.O.

3. What is the output of the program below, if you assume that the parameter passing mechanism is (a) call-by-value (b) call-by-reference (c) copy-restore (d) call-by-name? Show your rough work.

```
int a = 10;
void f(x, y, z) {  x = x + 1;    a = a * y;    z = z + a;  }
void main(void) {  int b = 5; f(a, a+b, b); printf("%d %d", a, b);  }
```

[3+2+4+2=11]

4. Consider the following function in C. Assume that (i) `data` is an array of `num` integers, and (ii) integers, floats and pointers occupy 4 bytes each.

```
float *mean_var(int *data, int num)
{ int i;
  float stats[2];
  /* body starts here */
  stats[0] = stats[1] = 0;
  for (i = 0; i < num; i++) {
      stats[0] += data[i];
      stats[1] += data[i] * data[i]
  }
  stats[0] /= num;
  stats[1] = stats[1]/num - stats[0] * stats[0];
  /* body ends here */
  return stats;
}
```

(a) Convert the body of the procedure into 3-address code. Each time a temporary variable is needed, use a new temporary. You may use the name of an array instead of the constant (base address) associated with that array. **Do not perform any optimization at this stage.**

(b) Assume that it takes 256 bytes to store the saved machine status. Draw a suitable layout for the Activation Record (AR), including byte offsets, for `mean_var`.

(c) Write the machine code for the calling sequence and return sequence for a call to `mean_var` from `main`. Assume that (i) the stack grows from low addresses to high addresses; (ii) the stack pointer points to the beginning (i.e. lowest address) of an AR; and (iii) the AR for `main` occupies 320 bytes.

(d) Identify the leaders in the 3-address code in (a) and draw the flow-graph.

(e) Optimize your intermediate code in (a) by using whichever of the following techniques are applicable: constant folding, global common sub-expression elimination, copy propagation, dead code elimination, code motion, induction variable elimination.

(f) Briefly explain the single most important flaw in the above code.

[15+6+(6+4)+(2+4)+8+2=47]

END

3. **(20 points)** Consider the sequential circuit as shown in Figure 2, in which two clocked D flip-flops are on feedback paths.

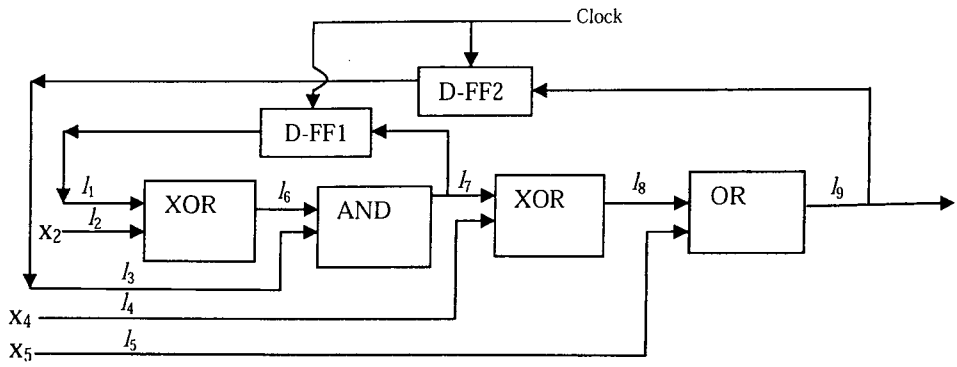


Figure 2. A sequential CUT

(a) Modify the above CUT for scan-based design by inserting a scan chain and show the associated control circuit.

(b) Determine a scan test vector for detecting l_8 stuck-at 0; show how the error can be observed by scan-in/scan-out mechanism.

(c) Assume both the FFs are initialized to 0. Show how a slow-rising (0 $\rightarrow$1) transition fault at line l_8 can be detected by applying a pair of scan-in vectors under launch-on-shift (LOS) scheme. Write the appropriate scan vectors along with primary input assignments and explain the procedure of observing the error, if any. (5+7+8)

4. **(15 points)**

(a) Let C be an n-input, single-output combinational circuit, which realizes a Boolean function F. Let l_i be an arbitrary line in C. Show that if C is irredundant, then a single stuck-at 0 (or a stuck-at 1) fault on line l_i can never produce a faulty function F_f which is equal to $\bar{F}$ (complement of F). (Hint: use the fact from Boolean difference that if a test set T_1 detects a stuck-at 0 and a test set T_2 detects a stuck-at 1 fault on the same line then they are disjoint, i.e., $T_1 \cap T_2 = \varnothing$.)

(b) An LFSR-based pseudorandom test pattern generator (TPG) has 5 parallel output lines y_1, y_2, y_3, y_4, y_5 that produce a sequence of 5-bit random test vectors, where in every vector, each bit has equal probabilities of 0's and 1's. A CUT is given which has 5 primary inputs. However, these 5 inputs require test vectors where 0's should appear with probabilities: {1/2, 3/4, 1/16, 5/8, 1/32}. Design the additional logic for producing such weighted random patterns. (8+7)

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination : (2012 - 2013)

Course Name : M.Tech (CS) **Year : 2nd year**

Subject Name : Neural Networks & Applications

Date : 21-02-13 **Maximum Marks : 50 Duration : 2 hrs**

Answer all the questions.

1. State and prove the fixed-increment convergence theorem for a single-layer perceptron. [5+20 = 25]

2. Show how a single-node perceptron model can be used to solve linearly separable two-class pattern classification problem. Consider n-dimesional patterns distributed in 2 classes. [15]

3. Describe the architecture and learning of a Radial Basis Function neural network. [10]

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination: 2012-2013 M.Tech. (CS) Second Year

Fault-tolerant Computing

Date: 21.2.2013 **Maximum marks = 60** Credit: 30%

Time: 3 hours

Name:_____ **Roll No.:**_____

Instructions (Read carefully)

A. This is an **OPEN BOOK/OPEN NOTES** exam. Answer all questions; partial credit may be given for incomplete/incorrect answers.

B. Total points = 70; **maximum score = 60.**

C. You may write your answer on the test booklet.

---

1. **(20 points)** Consider the following circuit-under-test (CUT) in Figure 1 with 5 Boolean inputs x_1, x_2, x_3, x_4, x_5, which consists of 2 XOR gates, one 2-input AND, one 2-input OR gate, producing a Boolean function F. The lines in the circuit are labeled as l_1, l_2, l_3, ..., l_9.

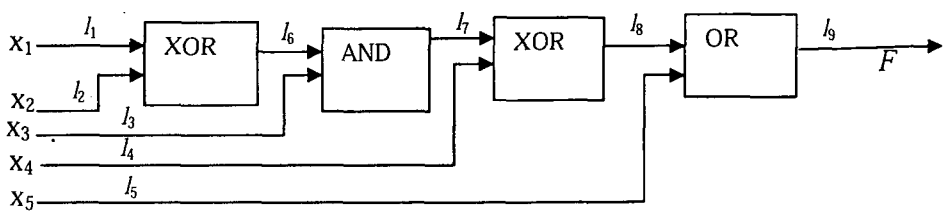


Figure 1. CUT

(a) Write down the output Boolean function F in terms of the primary input variables.

(b) Determine a test vector that detects the fault l_7 stuck-at 0.

(c) Ignore faults in the interior of XOR blocks and consider single stuck-at 0 and 1 faults only on lines l_1, l_2, l_3, ..., l_9. Write a complete test set T that gives 100% fault coverage. Justify your argument. (Hint: it is possible to find a test set T with only 5 vectors. If you can find such a set, show it; otherwise, you may determine another test set, but it should guarantee 100% fault coverage). (4+4+12)

2. **(15 points)** Referring to the CUT of Figure 1, please answer the following questions.

(a) List all the single stuck-at faults, which are detected by the test vector t: $\{x_1, x_2, x_3, x_4, x_5\} = 0\ 0\ 0\ 0\ 0$.

(b) Compute the Boolean difference (dF/dl_6) of the function F with respect to line l_6, and determine all the test vectors that are capable of detecting l_6 stuck-at 1.

(c) Does there exist a test vector t that detects l_2 stuck-at 0, which also detects l_5 stuck-at 0? Justify your argument. (6+6+3)

Indian Statistical Institute

Periodical Examination of Second Semester (2012-2013)
M. TECH. (CS) 2 Year
Subject: Information and Coding Theory
Date: 22/02/2013 Time: 2 hours Maximum Marks: 40

1. Prove that $I(X;Y) \geq 0$ with equality if and only if X and Y are independent. [5]

2. Prove that Huffman code is optimal. [10]

3. State the channel capacity theorem and give an outline of the proof. [10]

4. Let $\mathcal{X} = \{1,2,3\}$ and $X_1, X_2, X_3 \ldots$ be a time invariant Markov chain where X_is take values from the set $\mathcal{X}$. The transition matrix is

$$P = \begin{bmatrix} 2/3 & 1/6 & 1/6 \\ 1/2 & 0 & 1/2 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}.$$

Find the stationary distribution of the Markov chain. [5]

5. Let (X,Y) have the following joint distribution:

Y \ X	1	2	3	4
1	1/8	1/16	1/32	1/32
2	1/16	1/8	1/32	1/32
3	1/16	1/16	1/16	1/16
4	1/4	0	0	0

Find $H(X), H(Y), H(X|Y), H(Y|X)$ and $I(X;Y)$. [10]

INDIAN STATISTICAL INSTITUTE

Mid-semestral Examination:(2012-2013)

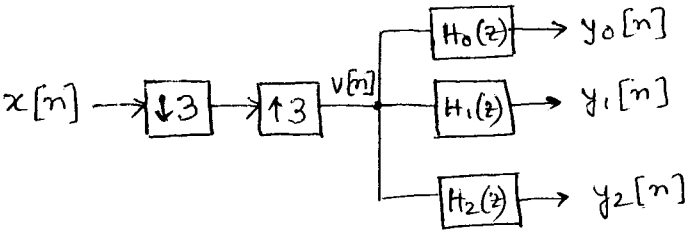
MTech C.S. 2nd Year

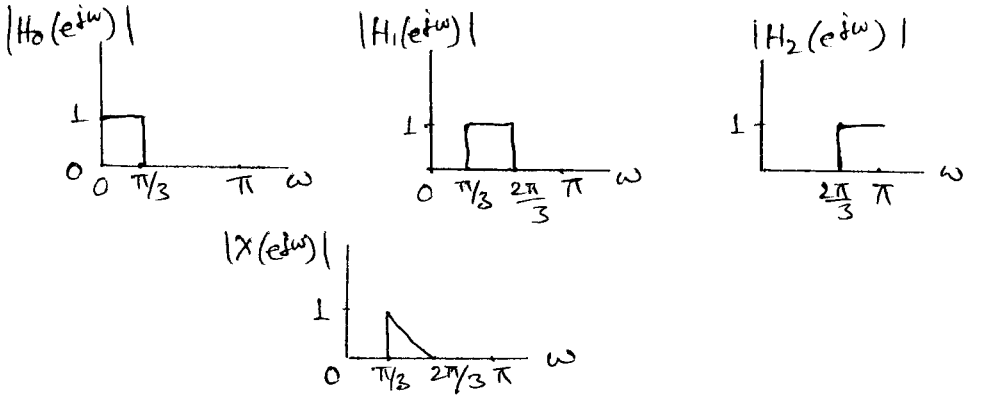
Advanced Digital Signal Processing

Date: 22.2.2013 Maximum Marks: 60 Duration: 2 hours

Note: The marks add up to 75. The maximum you can score is 60. The exam is open-book, open-notes. Use of calculators is permitted.

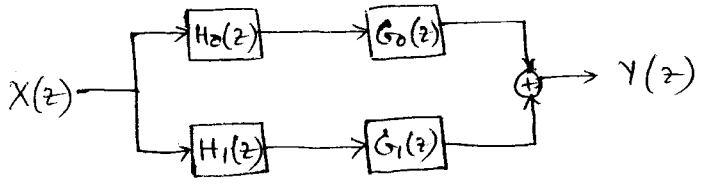
1. In the multirate structure shown below, $H_0(z)$, $H_1(z)$ and $H_2(z)$ are ideal, zero-phase, real-coefficient low-pass, band-pass and high-pass filters, with frequency responses as shown. The input is real, with a Discrete Time Fourier Transform (DTFT) as shown. Sketch the DTFTs of the outputs $v[n]$, $y_0[n]$, $y_1[n]$ and $y_2[n]$. [5+5+5+5]





2. Give a computationally efficient realization of a factor-of-5 interpolator using a length-15 linear phase FIR filter. [10]

3. In the analysis-synthesis filter bank shown below, $H_0(z) = \frac{1+z^{-1}}{1-0.2z^{-1}}$. Determine $H_1(z)$, $G_0(z)$ and $G_1(z)$ so that the filter bank performs perfect reconstruction. [7]



4. Consider the causal transfer function

$$H(z) = z^3 + 2.75z^2 + 1.3z + 0.3$$

Determine if all of its zeros are inside the unit circle. [10]

5. Show that

$$H(z) = \frac{0.07z^3 + 0.4z^2 + 0.73z + 1}{z^3 + 0.73z^2 + 0.4z + 0.07}$$

is an all-pass transfer function. Obtain the corresponding cascaded lattice structure. [5+10]

6. Design a two-stage system for reducing the sampling rate of a signal from 16 kHz. to 400 Hz. The specs for the decimation filter are : Passband edge: 180 Hz., Stopband edge 200 Hz. , Pass band ripple= stop band ripple=0.002. Use the Interpolated Finite Impulse Response (IFIR) approach for designing the filter. [13]

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination: (2012 – 2013)

M.Tech. (CS) II Year

Parallel Processing: Architectures and Algorithms

Date: Feb. 23, 2013 Duration: 3 hrs

NOTE: A student may answer all questions but maximum marks attainable is 100.

1. (a) According to the multiplicity of instruction and data streams, classify the parallel machine architectures with a schematic diagram for each.

   (b) Given two arrays A and B, each with n elements, consider the evaluation of the following inner-product expression:

   $$S = \sum_{i=1,n} A_i.B_i$$

   i) in an SISD machine with one PE, and
   ii) in an SIMD machine with m PEs connected by a bidirectional circular ring, $m < n$.

   It is assumed that each ADD operation requires two time units and each MULTIPLY operation requires four time units. Data shifting along the circular ring between adjacent PE's requires one time unit. Develop an algorithm for computing S on the SIMD machine mentioning the initial data distribution. Hence calculate the speed up of using the SIMD machine over the SISD machine for the evaluation of S.

   [4 + (10 + 6) =20]

2. How many steps are required at maximum to broadcast a single data item from one PE to all other PE's in each of the following interconnection networks with N nodes, $N= 2^{2n}$:
   i) A $2^n \times 2^n$ mesh network
   ii) A hypercube network.

   Justify your answer.

   [8+8=16]

3. (a) Draw the block diagram of an 8×8 omega network using 2×2 switches. Show that it is a *full-access unique-path* multi-stage interconnection network.

   (b) Given a BPC permutation P: $x_2\, x_1\, x_0 \rightarrow x_0\, x_2'\, x_1$, draw the conflict graph and find the minimum number of passes required to route P. Is P admissible in the omega network with one or two additional stages? Justify your answer.

   [(4+5)+(8+6)=23]

   [PTO]

4. (a) Given two processes P_i and P_j with corresponding sets of inputs and outputs as I_i, I_j and O_i, O_j respectively, state the conditions to be satisfied for P_i and P_j to be executable in parallel, i.e., *parallelizable*. Is the relation of *parallelizability* an *equivalence relation*? Justify your answer.

(b) Consider the following program segment with six instructions:

$P_1 : A = B + C$

$P_2 : D = B \times C$

$P_3 : E = B \times A$

$P_4 : F = D + C$

$P_5 : B = E + A$

$P_6 : C = D + A$

Draw the data dependence graph considering each statement as a process.
Show an efficient scheduling of the processes exploiting the maximum parallelism existing among the processes. Ignore the resource dependences. Assuming that addition takes 10 time units, multiplication 100 time units, and inter-processor communication 300 time units, calculate the speed-up achieved by your scheduling.

[(3+3)+(7+5+6) = 24]

5. (a) Define a Cayley graph. Prove that a Cayley graph is always vertex symmetric.

(b) Define an *n*-star graph and find its diameter.

[(3+5)+(4+8) = 20]

6. (a) In a regular graph with N nodes each of degree d, find out an expression for the lower bound on the diameter k. What is Moore graph?

(b) In an $N \times N$ mesh network, how many distinct shortest paths exist between a pair of nodes (x_1, y_1) and (x_2, y_2), $0 \leq x_1, y_1, x_2, y_2 < N$, where (x_1, y_1) is the node at row x_1 and column y_1. Two paths are distinct if they differ at least in one edge. How many paths of these are node disjoint (no common node, except the source and destination)?

[(5+3)+(8+4) = 20]

7. Write a CUDA program to search an element x within a given array A with 2^{20} elements.

[15]

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination

M.Tech(CS)-II year (Semester II)

Computational Geometry and Multidimensional Search

Date: 25th February, 2013 Maximum marks:60 Time: 2:30 Hrs.

Attempt as many questions as you wish, but maximum marks you can score is 60.

1. Let P be a simple (non-convex) polygon. Describe an efficient algorithm that computes the convex hull of P. [10]

2. Let S be a set of n circles in the plane. If k is the number of intersection points between the circles, then describe an efficient algorithm in terms of n and k that computes all intersection points. [15]

3. Let S_1 be a set of n disjoint horizontal line segments and let S_2 be a set of m disjoint line segments. Give an efficient algorithm to count how many intersections there are in $S_1 \cup S_2$. [15]

4. Let $C_1, C_2, C_3, \ldots C_n$ be n convex sets in the plane such that intersection of every triple contains a line segment of length 1. Prove or disprove whether the intersection of all the convex set contains a line segment of length 1. [20]

5. Let $P_1, P_2, P_3, \ldots P_m$ be m convex polygons not necessarily disjoint in the plane with n vertices in total such that no vertex is common to two or more polygons and vertices form a point set in general position (that is no three points are collinear). Find in terms of complexity, the number of lines that intersects all the polygons and are tangent to atleast two of them. [Hints: there may be some relation with maximum length of a Devenport-Schinzel sequence of some order.] [20]

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination: 2012-2013

M. Tech. (CS) II year

Data Mining and Knowledge Discovery

Date: 25/02/13 Maximum Marks: 50 Duration: 2 hours

1. (i) What are some characteristics of a good clustering algorithm in the context of data mining?
   (ii) What do you understand by the "scalability" of an algorithm?
   (iii) Outline algorithm PAM, commenting on its merits and demerits.

   [4+4+8]

2. (i) Mention some of the similarity/dissimilarity measures, used in clustering, for different types of data.
   (ii) What are the major clustering approaches? Mention a few examples of each.
   (iii) How is cluster similarity defined in hierarchical clustering?

   [5+7+6]

3. (i) Describe algorithm CURE.
   (ii) In what ways does it perform better than some existing approaches, and why?
   (iii) How do we validate a clustering result?

   [8+4+4]

INDIAN STATISTICAL INSTITUTE
M. Tech (CS) II year : 2012–2013
Quantum Information Processing & Quantum Computation
Periodical Examination

Date: 26. 02. 2013 Maximum Marks: 50 Time: $2\frac{1}{2}$ Hours

Answer any part of any question. Maximum marks you can obtain is 50. The paper is of 60 marks.

Please answer all parts of a question at the same place.

1. Explain, with examples, how
   (a) an n-qubit state and
   (b) an n-qubit input, n-qubit output quantum gate

   can be represented in matrix form. 5+5 = 10

2. Explain how quantum entanglement can be exploited in
   (a) super dense coding,
   (b) teleportation and
   (c) remote state preparation.

   3+4+3 = 10

3. (a) Explain the problem that is solved by the Deutsch-Jozsa algorithm.
   (b) Describe the algorithm clearly.
   (c) Consider that a Boolean function $f(x_1, x_2, x_3) = x_1 \oplus x_2 \oplus x_3$ is used in the Deutsch-Jozsa algorithm as unitary transform U_f. What will be the output observed?
   (d) What is the importance of this algorithm in contrast to classical paradigm?

   2+8+8+2 = 20

4. (a) Explain the problem that is solved by the Grover algorithm.
   (b) Describe the algorithm clearly with all the related complexity issues.

   5+15 = 20

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination: 2012-13(Second Semester)

Course Name: M.Tech. (CS) 2nd Year

Subject Name: Natural Language Processing

Date: 26.02.2013 Maximum Marks: 50 Duration: 2 hours

Note: Answer all the questions.

Q1. Consider an English example:

Which book did she review without reading?

Here *reading* has a "gap" as shown below with underscores:

Which book did she review without reading____?

We don't know "without reading what?". This gap is considered as "parasitic" gap (parasitic on "review what?") and cannot easily stand on its own.

Assume that "parasitic gap" is really a rare syntactic construction and this occurs on average once in 100,000 sentences. You have developed an approach to identify sentences with parasitic gaps. Your approach is pretty good, but not perfect. If a sentence has a parasitic gap, your method will say so with probability 0.95 and if it doesn't, it will wrongly say it does with probability 0.005. Suppose the test says that a sentence that contains a parasitic gap. What is the probability that this is true? Explain your answer.

[8]

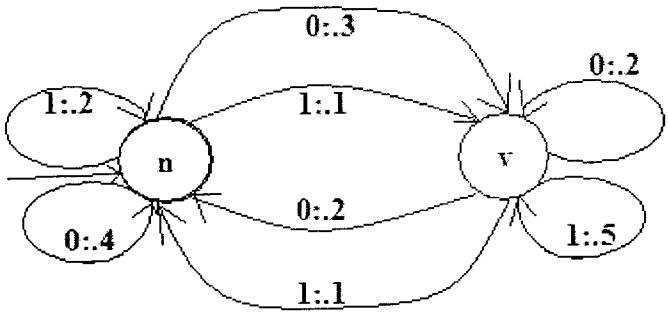
Q2. Consider a dictionary having the following root words: *bird, cat, dog, fish, frog*. Each word can be used in singular and plural form. The plural forms of the words are *birds, cats, dogs, fish, frogs*. The words have their Babytalk versions too: *birdie* (from *bird*), *fishie* (from *fish*), *doggie* (from *dog*), *froggie* (from *frog*), *kittie* (from *cat*). The plural forms of Babytalk words are: *birdies, fishies, doggies, froggies, kitties*.

Develop a model for building a morphological analyzer for the above morphological data. Given a word, the analyzer will return its corresponding root word, singular/plural, Babytalk (yes/no). For example, when the word *dog* is given, your analyzer returns: ROOTFORM: dog, NUM: singular, BABYTALK: false. When the word doggies is given, the result is: ROOTFORM: dog, NUM: plural, BABYTALK: true.

Develop and explain only the model that you like to propose for the above problem. No coding/implementation details need not be given or discussed.

[15]

Q3. Consider a two-state simple Hidden Markov Model as given below.



Compute the best state sequence when the input is "1110".

[11]

Q4.

(a) Though there are many well-established techniques for parsing a context-free grammar, explain why one particular parsing scheme, chart parsing, is very common in natural language processing. Compare the time and space complexity of chart parsing (Earley Algorithm) with that of CYK.

[5]

(b) Consider the following context-free grammar:

S	→	NP VP				
VP	→	V NP	|	VP PP	|	V
NP	→	NP PP	|	N		
PP	→	PP NP	|	Prep NP		
N	→	*salesperson*	|	*dog*	|	*biscuits*
V	→	*sold*				
Prep	→	*the*				

Show the intermediate steps if Earley Algorithm is followed to parse the sentence *salesperson sold the dog biscuits* following the CFG given above. In order to get more space (breadth-wise), you may show the chart in the landscape mode in your answer script.

[11]

INDIAN STATISTICAL INSTITUTE

Mid Semester Examination

M. Tech (CS) - II Year, 2012-2013 (Semester - IV)

Computational Complexity

Date : 26.02.2013 Maximum Marks : 60 Duration : 2.5 hours

Note: Answer as much as you can, but the maximum you can score is 60.
Write your answers in the same order as that of the questions given below.

(Q1) A deterministic 2-dimensional Turing machine is like a standard Turing machine except that each of its tapes is a two-dimensional grid with a top end and a left end but extends infinitely to the right and down.. Each tape head, is initially at origin $(0,0)$ and can move one tape cell not only $left$ or $right$, but also up or $down$. The input string x of length n is written in tape cells $(0,1),(0,2),\ldots,(0,n)$, the origin $(0,0)$ having a special symbol O, and all other cells (i,j) has a special blank symbol b. It accepts the input by erasing the entire board, i.e., filling all cells with b.

(i) Give a formal definition of such a 2-dimensional TM.

(ii) Argue (semi-formally) that this 2-dimensional TM can be simulated by a standard TM. Further, show that for every time-constructible function $T : \mathbb{N} \to \mathbb{N}$, and every Boolean function f, if f can be computed by a 2-dimensional TM in time $T(n)$, then $f \in \mathbf{DTIME}(T(n^2))$.

$[6 + 10 = 16]$

(Q2) (a) Show that the set

$\{M | M$ is a Turing machine which halts on all inputs of length less than 2013$\}$

is recursively enumerable but its complement is not.

(b) Show that the integer square root function is primitive recursive. On input n, the function should return the greatest integer less than or equal to the square root of n.

$[(4 + 3) + 7 = 14]$

(Q3) (a) Let $Double\text{-}SAT = \{\phi \mid \phi$ is a Boolean formula and has at least two satisfying assignments$\}$. Prove that $Double\text{-}SAT$ is NP-complete.

(b) Show that if $P = NP$, then we can factor integers in polynomial time.

[7+8=15]

(Q4) (a) Show that log-space reduction $\leq_L$ is transitive.

(b) Show that $A = \{w | w = w^R\} \in L$ but not is $SPACE(s)$ for $s = o(\log n)$.

(c) Give an example of a recursive function which is not space-constructible.

[6 + 6 + 3 = 15]

(Q5) (a) (i) Define the configuration graph $G_{M,x}$ of a non-deterministic Turing machine M on some input of length n. (ii) If M is a space-$S(n)$ machine, then determine the upper bound on the number of nodes of $G_{M,x}$.
(iii) Prove that there is an $O(S(n))$-size CNF formula $\phi_{M,x}$ such that for every two strings C, C' $\phi_{M,x}(C, C')$ = 1 if and only if C and C' encode two neighbouring configurations in $G_{M,x}$.

(b) If $PSPACE = \bigcup_i \Sigma_i^P$, then show that for some i, $\Sigma_i^P = \Sigma_{i+1}^P$.

[(4+3+7)+6=20]

INDIAN STATISTICAL INSTITUTE

Mid-Semestral Examination

M.Tech(CS)-II Year, 2012-2013 (Semester-IV)

Distributed Computing

Date: February 27, 2013 Maximum Marks: 70 Duration: **3 Hours**

Note: The paper carries a total of 85 marks. You may answer as much as you can. The maximum you can score is 70.

1. (a) Describe the conditions to realize the *happened before* relation by using Lamport's logical clocks. What are the limitations of Lamport's logical clock?
   (b) Describe the Huang's Termination detection algorithm. Prove the correctness of the algorithm.
   (c) Describe Birman-Schiper-Stephenson's protocol for causal ordering of Messages.

   ((3+2)+(6+4)+5=20)

2. (a) Show that in Lamport's algorithm for distributed mutual exclusion if a site S_i is executing the critical section, then S_i's request need not be at the top of the *request-queue* at another site S_j. Is it still true when there are no messages in transit?
   (b) Describe the Suzuki-Kasami's broadcast algorithm for distributed mutual exclusion. Prove the correctness of the algorithm. Comment on the message complexity and synchronization delay of the algorithm.
   (c) What are the drawbacks of Maekawa's algorithm for distributed mutual exclusion. How would you resolve these drawbacks?

   ((3+2)+(7+3+2)+8=25)

3. (a) Analyze the truth of the following claim with justification/counter example:
   "In the OR-model, a deadlocked process must belong to one or more directed cycles in the wait-for graph."
   (b) The k-exclusion problem is similar to the mutual exclusion problem, but here at most k nodes can be in the critical section at the same time, for some predefined constant k. Design an algorithm to achieve k-exclusion in a distributed system. Analyze the message complexity (per critical section entry) of your algorithm.
   (c) Describe Chandy-Misra-Haas's distributed deadlock detection algorithm for the AND request model.
   (d) Give an example to show that in the AND request model, false deadlocks can occur due to deadlock resolution in distributed systems.

   (6+4+6+4=20)

4. (a) A ***fragment*** is a subtree of an MST (minimal spanning tree). If F is a fragment of an MST and e is the least-cost (unique) outgoing edge of F, then prove that $F \cup e$ is a fragment.
   (b) In GHS algorithm for MST, describe the combining strategies for fragments. Prove that if these rules are obeyed, the number of times a process changes its fragment name or level is at most $N \log N$ where N is the number of processes.
   (c) Prove that in GHS algorithm for MST no deadlock cycle can occur.

   (4+5+5+6=20)

Indian Statistical Institute

Advanced Image Processing

M.Tech.(CS)-II Year: 2012-13

Full marks: 60 Time: 2 Hours

Date: 27.02.2013

Answer **any six** questions. All questions carry equal marks.

1. Assuming pin-hole camera model for perspective projection from 3-D to 2-D, prove that
   (i) a straight line in 3D is mapped to a straight line in 2D, and
   (ii) Define vanishing point. Derive its coordinate. [5+5]

2. a) State three basic principles of photometric model of image formation.
   b) Derive the equation (all terms have their usual meaning):

$$g(x,y) = \int_{\infty}^{\infty} \int_{\infty}^{\infty} h(x-\alpha, y-\beta) f(\alpha,\beta) d\alpha d\beta + \eta(x,y)$$

   [3+7]

3. Derive the expression of Wiener filter for image restoration using minimum mean-square estimation approach. [10]

4. a) Define image gradient following best plane fit approach.
   b) Prove that: best plane fit gradient $= \sqrt{2}$ Roberts gradient [8+2]

5. a) Prove that the Fourier transform of a two-variable function $f(x,y)$ is rotated by an angle θ if $f(x,y)$ is rotated by the same angle.
   b) Obtain the Fourier transform of the Laplacian of a two-variable function $f(x,y)$. Assume that x and y are continuous variables. [6 + 4]

6. Describe the Fast Fourier Transform algorithm and discuss its computational complexity. [8 + 2]

7. a) State and prove the convolution theorem.

b) Prove that the origin of the Fourier transform of a two-variable function $f(x, y)$ can be moved to the center of its corresponding $N \times N$ frequency square by multiplying $f(x, y)$ by $(-1)^{x+y}$. [7 + 3]

8. a) Calculate the sequency of each column of Hadamard matrix of order 8.

b) Describe the Hotelling transform and show that it is optimal in the least-square-error sense. [3+(5+2)=10]

INDIAN STATISTICAL INSTITUTE

Mid Semestral Examination

M. Tech (CS) - II Year, 2012-2013 (Semester - IV)

Topics in Algorithms and Complexity

Date : 28.02.2013 Maximum Marks : 60 Duration : 2.5 Hours

Note: Answer as much as you can, but the maximum you can score is 60.

This is a 2 page question paper with 7 questions.

(Q1) Given an undirected graph $G = (V, E)$ with $|V| = n$ vertices and $|E| = m$ edges. Show that there exists a *partition* of V into two disjoint sets V_1 and V_2 such that at least $m/2$ edges cross the partition. [10]

(Q2) Consider the randomized algorithm for the two dimensional linear programming studied in class.

(a) Extend the algorithm so that it works for d dimensions, where $d \geq 2$.

(b) Deduce the time complexity as a function of n and d, where n is the number of linear constraints.

[5 + 6 = 11]

(Q3) Consider an instance of SAT with m clauses, where every clause has exactly k literals. Design a Las Vegas randomized algorithm that finds an assignment satisfying at least $m(1 - 2^{-k})$ clauses. Analyze the expected running time of the algorithm. [10 + 5 = 15]

(Q4) Consider the following randomized version of quick sort.

Algorithm 1: Randomized quick sort algorithm.

Input: An array $A[1 \ldots n]$ of n unique numbers.

Output: The elements of A sorted in increasing order.

1 Choose an element a uniformly at random from A;

2 Compare each element of A with a and determine the set A_1 of all numbers less than a and the set A_2 of all numbers greater than a ;

3 Recursively sort A_1 and A_2;

4 Output the sorted version of A_1, followed by a, followed by A_2;

Analyze the expected number of comparisons the above algorithm does. [12]

(Q5) Let $G = (V, E)$ be a graph where we assign any of the three colors {R, B, G} to vertices in G. We say an edge (u, v) is *conflict free* if the vertices u and v are assigned different colors. Consider a coloring that maximizes the number of *conflict free* edges and let this number be c^*. Notice that this is an optimization version of the 3-COLORING problem's decision version which is NP-hard.

Consider the following randomized polynomial time algorithm to solve the above problem approximately. We pick any one of the three colors uniformly at random and color a vertex. We do this for all vertices in G.

Using the above algorithm, what will be the expected number of edges that would be *conflict free*? Explain your result. [10]

(Q6) Suppose we roll a standard fair die 200 times. Let X be the sum of the numbers that appear over the 200 rolls. Use Chebyshev's inequality to bound $\mathbf{Pr}[X \geq 750]$. [5]

(Q7) Suppose that we have an algorithm that takes as input a string of n bits. We know that the expected running time is $O(n^2)$ if the input bits are chosen independently and uniformly at random. What can Markov's inequality tell us about the worst-case running time of this algorithm on inputs of size n? [12]

INDIAN STATISTICAL INSTITUTE

M. Tech. (CS) II Year (2012-13), II semester

Periodical Examination

ADVANCED PATTERN RECOGNITION

Date: March 2, 2013 Duration: 150 minutes Marks: 50

Note: Answer all the questions.

1. Suppose you have two binary strings A = 00101 and B = 10001. Let the crossover probability be p. Find the probability that, after crossover operation, the resultant two strings will be A and B. [5]

2. Describe a data condensation procedure for data mining. [5]

3. (a) Describe Parzen's density estimation procedure.
   (b) Describe its generalization to multivariate case. [5+5=10]

4. Describe the Perceptron algorithm for classification. State also the assumptions under which it converges. [10]

5. (a) Describe the k-nearest neighbor density estimation procedure, and derive the k-nearest neighbor rule using the density estimation procedure.
   (b) Describe an algorithm for reducing the size of the training sample set for k-nn decision rule. [(5+5)+5=15]

6. Describe a density based clustering algorithm. [5]

INDIAN STATISTICAL INSTITUTE

Mid Semester Examination

M. Tech (CS) - II Year, 2012-2013 (Semester - IV)

VLSI Design Algorithms

Date : 04.03.2013 Maximum Marks : 50 Duration : 2.5 hours

Note: Answer as much as you can, but the maximum you can score is 50.
Write your answers in the same order as that of the questions given below.

(Q1) (a) Consider the following netlist NL_1. Assume that the first gate is the source, and area of all the eight gates are 1.
NL_1: n_1= *a-b-c*, n_2 = *b-d-e-f*, n_3 = *c-f-g*, n_4 = *a-g*, n_5 = *d-e-h*, n_6 = *f-h*.
(i) Model the netlist with an edge-weighted undirected graph G_1. The edge weight is 1/(k-1) where k is the number of gates, e.g., 1/2 for n_1 and 1/3 for n_2.
(ii) Given an initial partition $P_1 = \{(a, b, c, d), (e, f, g, h)\}$ of G_1, perform a single pass of Kernighan-Lin algorithm. Give the cutsize after each swap. What are the initial, final, and best cutsizes?

(b) (i) Model NL_1 with a non-weighted hypergraph H_1.
(ii) Given an initial partition P_1 (from 1(a)(ii)) of H_1, calculate the initial gains of each gate and setup the two gain-buckets in the Fiduccia-Mattheyeses algorithm.

$[(3 + 6) + (2 + 4) = 15]$

(Q2) Consider the following slicing floorplan with 8 modules represented by a reverse Polish expression: 235V4HV18VH7V6V. The dimensions of the modules 1 through 8 are $\{(2, 4), (2, 4), (3, 3), (5, 3), (1, 3), (1, 4), (3, 6), (4, 2)\}$ respectively.

(a) Draw the corresponding slicing tree.

(b) Compute the area of the smallest rectangle that can accommodate the modules with no overlap. Each module is free to rotate by 90°.

(c) Draw the optimally-sized slicing floorplan using (x, y) coordinates.

$[4 + 7 + 4 = 15]$

(Q3) (a) Formulate an ILP-based routing problem for the following netlist: $\{n_1 = B_1 - B_4 - B_6, n_2 = B_3 - B_4, n_3 = B_1 - B_2 - B_5 - B_6, n_4 = B_2 - B_5\}$ The placement of the blocks are as shown below. Assume that the capacity of each boundary is as follows:
$\{B_1 - B_2 : 3, B_2 - B_3 : 4, B_4 - B_5 : 5, B_5 - B_6 : 4, B_1 - B_4 : 7, B_2 - B_5 : 5, B - 3 - B_6 : 3\}$.

B1	B2	B3
B4	B5	B6

(b) Explain briefly the Akers' coding scheme for maze routing. How is it helpful?

[10+5=15]

(Q4) Consider the following problem instance for two-layer channel routing. Assign tracks from top to bottom and left to right.
TOP=[1,4,1,0,4,3,2,5,6,5], BOT=[2,1,0,4,5,3,6,6,3,6]

(a) Construct the HCG & VCG and compute the minimum channel density.

(b) Apply the unconstrained Left Edge (LE) algorithm.

(c) Apply the constrained LE algorithm.

(d) Decompose multi-terminal nets into two-terminal nets at terminal positions and construct its VCG.

(e) Apply the constrained LE algorithm to part d.

(f) Construct the zone representation for the Yoshimura-Kuh algorithm.

$[6 + 4 + 3 + 3 + 4 = 20]$

(Q5) (a) Why does the clock net need special routing algorithm?

(b) Illustrate the construction of the global routing graph and the detailed routing graph with an example of a given FPGA placement having four CLBS and two 3-terminal nets.

(c) Give the main categories of algorithms for standard cell placement with an example for each category. $[2 + (5 + 5) + 3 = 15]$

Indian Statistical Institute

Mid-Semester Examination 2012-13
M. TECH.(CS) II Year
Subject: Document Processing and Retrieval
Full Marks: 50 Duration: 2 hrs.

(Answer all questions)

1. What is a skew of a document image? What are the differences between tilt and skew? Describe a Hough Transform based method to detect the skew of a document image. [2+3+7]

2. Generate a feature based tree classifier for the recognition of the following printed alphanumeric characters (P, R, 3, S, Z, 2, A, K, 6, 9, c, e). [12]

3. Draw a basic block diagram of a document analysis system. Discuss about OCR technology for multi-script documents. Explain different types of character touching that may appear in a document image. [4+3+7]

4. Discuss a method to separate graphics and text from a document page. [6]

5. Let $x_1, x_2, x_3 \ldots\ldots x_n$ be n consecutive contour pixels of a character. Discuss a rotation invariant feature extraction method based on $x_1, x_2, x_3 \ldots\ldots x_n$. [6]

INDIAN STATISTICAL INSTITUTE

Mid-Semester Examination : (2012-2013)

M.Tech.(CS) II Year

Advanced Database Theory & Applications

Date: 05.03.2013 **Maximum Marks: 60** **Duration: 2.5 Hours**

1. Three transactions T1, T2 and T3 are executed concurrently according to the following schedule sharing two different data items X and Y:
   T1: R(X), T3: R(Y), T2: W(X), T3: W(Y), T3: R(X), T1: R(Y), T2: W(Y).
   Where, R(Q) = read data item Q.
   W(Q) = write data item Q.
   a) Drawing a precedence graph show whether the above schedule is conflict serializable?
   b) To examine view serializability, draw the labeled precedence graph for two data items separately to show whether they are individually view serializable?
   Also draw the composite labeled precedence graph to examine whether the schedule is view serializable considering both the data items together?
   (5+5+5+4=19)
2. Consider that the schedule given in Question 1 is executed using timestamp ordering protocol where a transaction is rolled back for any time conflict. If TS(T) signifies the timestamp of any transaction T, timestamp ordering of the three transactions given in Question 1 is,
   $$TS(T1) > TS(T2) > TS(T3)$$
   What would be the status of the three transactions at the end of the schedule and what would be the value of the read and write timestamps of the different data items. Consider the initial value of all the read and write timestamps to be less than TS(T3).
   (7x3=21)
3. Considering the concurrent schedule given in Question 1, examine whether the schedule is executable under two-phase locking protocol with upgrade facility?
   (6)
4. Let T1, T2, T3 and T4 are four transactions accessing a common data item. TS(T) represnts the timestamp of the transaction T. The timestamps of the four transactions are related as, $TS(T1) < TS(T2) < TS(T3) < TS(T4)$.
   In order to access the common data item, the transactions arrive and form a queue in the order (T2, T1, T4, T3). The main memory can accommodate two transactions at a time for concurrent execution. However, the transactions are picked up sequentially from the queue and resources are allocated to each transaction in the order of its appearance in the queue. Since the transactions are sharing a common data item, it is assumed that the execution of one transaction will always hinder the execution of any other. Once a transaction is chosen for execution, all its required resources are allocated, if available. Otherwise, the transaction is either rolledback or made to wait depending on the deadlock prevention scheme in use. Any transaction once rolledback is re-scheduled immediately and it joins at the end of the queue with its timestamp unaltered. Count the number of rollbacks that would occur before all the four transactions are completed, if the system employs,
   a) Wait-die scheme for deadlock prevention
   b) Wound-wait scheme for deadlock prevention.
   (7x2=14)

※

INDIAN STATISTICAL INSTITUTE
End-Semester Examination: 2012-13
Course Name: M.Tech. In Computer Science
Subject Name: VLSI Design Algorithms

Date: 22. 04. 2013 *Maximum Marks: 100* *Duration: 3 hours*

Answer any five questions.

---

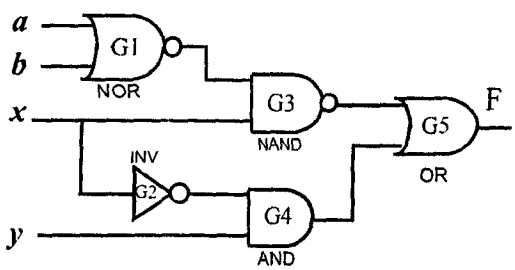
1. [5 X 4 = 20]

(a) Consider the following version of the Shannon expansion theorem.

$$F(x_1,\ldots x_i,\ldots x_n) = x_i \cdot F(x_i = 1) + \sim x_i \cdot F(x_i = 0)$$

The above can be thought of as a "sum of products" form, since the equation is an OR (sum) of two small AND (product) terms. But there must be a "*product of sums*" form for the Shannon expansion. Use Boolean algebra, and the basic properties of cofactors, and deduce a "*product of sums*" form of the Shannon expansion.

(b) Consider the small logic network shown below. Use the definition of the Boolean difference and compute the values of $\partial F / \partial x$ and $\partial F / \partial y$.



(c) Considering the same gate network as above, which are correct Boolean formulas for these quantification operations on F?

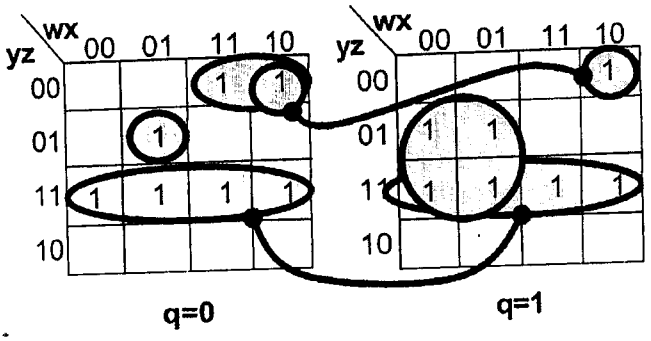
i) **$(\forall x, F)[a, b, y] = a + b$**

ii) **$(\exists x, F)[a, b, y] = \sim a + \sim b + \sim y$**

iii) **$(\exists x, F)[a, b, y]$ is a constant (0 or 1)**

iv) **$(\forall x, F)[a, b, y] = (\sim a + \sim b).\, y$**

(d) Consider the 5-variable Karnaugh map shown below, which shows a function of variables q,w,x,y,z. Karnaugh map groupings (covers) that span both sides of the Kmap (q=0 and q=1) are linked with a solid line with small circles at each end.



Looking at the cover of this function, select which of the following are correct statements.

i) This is a prime cover of this function.

ii) If we translate this cover into a standard SOP form with AND gates and an OR gate, that logic will have 18 literals.

iii) If we translate this cover into a standard SOP form with AND gates and an OR gate, that logic will have 10 literals.

iv) This cover has 5 cubes in it.

[8 + 3 X 4 = 20]

2.

(a) What do you mean by technology mapping? Clearly explain the tree covering algorithm for technology mapping with the help of an example.

(b) Answer the following with reference to the language Verilog.

i) What are the main differences between the "always" and the "initial" blocks?

ii) An "always" block always specifies a sequential circuit. Justify or contradict.

iii) What are the differences between blocking and non-blocking assignments within an "always" block?

iv) Event driven simulations cannot simulate sequential circuits.

[10 + 10 = 20]

3.

(a)

(i) Give a graph theoretic formulation of the 2-layer constrained via minimization problem.

(ii) Comment on its worst case time complexity

(b) Define the over-the-cell routing problem. Present an algorithm to solve it.

[(2 X (1+2) + 4 + 5 + 5 = 20]

4.

(a) For the following types of floorplans, give an example in each case and an appropriate representation scheme

(i) mosaic

(ii) compacted

(b) Give an example or counterexample, (as the case maybe), for the following statement:

"Soukup's maze router always produces the shortest path between two given terminals".

(c) Consider the following circuit code for simulation:

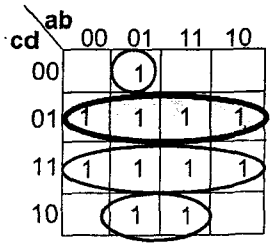
```
//circuit being simulated
initial
begin
 clk = 1'b0;
 #1 clk = ~clk;
 #1 clk = ~clk;
 #1 finish;
end
always @(clk)
begin
 a = b & c;
 if (a == 1'b0)
   p = q <<3;
end
```

Show the generated code for the above for both the interpreted and compiled code simulators.

5. [10 + 10 = 20]

(a) A 16-bit parity generator is to be designed using combinational logic. Compare the merits and demerits of two-level versus multi-level realization for the same.

(b) Consider a function F(a,b,c,d) which we are optimizing using the reduce-expand-irredundant method like ESPRESSO. Assume we have just done a REDUCE step and we have this intermediate, nonprime 4-cube cover of F:



Perform an EXPAND operation on the highlighted cube: **~cd**. Clearly show each of the following steps which ESPRESSO would follow to do this: build a cover of the OFF-set for this current cover, build a Blocking Matrix for the cube we seek to expand, and then computing a cover of this matrix.

6. [10 + 10 = 20]

(a) Design a circuit in structural Verilog (using Verilog primitive gates only), which will check the following relational operator:

A <= B, where both A & B are vectors of size 4.

Note: The module will have two output ports, which may be named out1 and out2. HIGH value in out1 will denote A < B and the same in out2 will denote A==B.

(b) Using the module created in Step 1 as primitive, design a 32 bit magnitude comparator that will take two 32 bit numbers and will specify the output through 3 output ports of one-bit each, which may be named LESSER, EQUAL, and GREATER.

INDIAN STATISTICAL INSTITUTE

Semestral Examination

M. Tech. - II Year (Semester - II)

Multi-dimensional Search and Computational Geometry

Date : 22.04.13 Maximum Marks : 100 Duration : 3:30 Hours

Note : You may answer any part of any question, but maximum you can score is 100.

1. You are given a set of n points P in the plane and an integer k, where $3 \leq k \leq n$. Define a k-corridor to be the closed region of the plane bounded by two parallel lines that contains at least k points of P. The height of a corridor is the vertical distance h between these lines.

   Present an $O(n^2 \log n)$ time algorithm for computing the minimum height k-corridor, where P and k are given as inputs. [20]

2. You are given a collection of n lines $L = \{\ell_1, \ell_2, \ldots, \ell_n\}$ in the plane, where, for $1 \leq i \leq n$, ℓ_i is given by the equation $\ell_i : y = a_i x + b_i$. You are also given two constants u and v. Give an $O(n \log n)$ time algorithm that counts the number of intersections of these lines that lie within the vertical lines $y = u$ and $y = v$. [20]

3. You are given a convex polytope with n vertices in $\Re^3$. Write an algorithm to preprocess them and describe a data structure such that for a query point q, testing whether the point q lies inside can be reported in order of polylog n time. [20]

4. You are given a collection of m pairwise disjoint x-monotone polygonal chains that extend from $x = -\infty$ to $x = +\infty$. Let n denote the total number of vertices in all these chains. Write an efficient algorithm for preprocessing these chains and for building a data structure so that, given any query point q in the plane, it is possible to report all chain edges that are intersected by a ray shot vertically downwards from q. The query time should be $O(k+\log n)$, where k is the number of edges hit by the ray. [20]

5. Let P be a convex polygon with n edges in the plane, which contains the origin in its interior. Consider each edge of P as a site, and let $Vor(P)$ denote the Voronoi diagram of these n line segments, restricted to the interior of P.

   What is the shape of each Voronoi edge and cell of the diagram? What is the combinatorial complexity of the diagram? Justify your answes. [10+10=20]

6. The weight of a triangulation is the sum of the length of all edges of the triangulation. Prove or disprove whether the Delaunay triangulation is a minimum weight triangulation. [20]

7. In a Voronoi diagram of a set of sites $P = \{p_1, p_2, \ldots, p_n\}$ in the plane, you observe that the Voronoi edge between sites p_i and p_j is semi-infinite (that is, one end of the edge goes to infinity). What can be said about the relationship between these two sites and the rest of the point set? Explain. [10]

Indian Statistical Institute

Advanced Image Processing

M.Tech.(CS)-II Year, 2012-2013

Full marks: 100 Time: 3 Hours

Date: 22.04.2013

Answer any **ten** questions. All questions carry equal marks.

1. a) Define dilation and erosion.

   b) Prove that dilation and erosion are dual operations.

   c) Assuming 4-connectivity of objects, write down the expression along with structuring elements for computing Genus or Euler number. [2+4+4=10]

2. a) Define moment and central moment in respect of an image.

   b) Derive the expression of *Principal Axis* of an object in an image. [4+6=10]

3. a) Define 'bpp' and 'PSNR' in respect of image compression. What is meant by image compression in terms of 'bpp' and'PSNR'?

   b) How is intensity of an image block represented using Block Truncation Coding (BTC)?

   c) Calculate the bpp of a 4x4 block of 256-level gray image compressed using BTC. [3+6+1=10]

4. a) What are active contours? Why are they used?

   b) Describe internal and external energy functionals in image forces. [4+6=10]

5. a) What are Von Neumann Neighborhood and Moore Neighborhood with respect to cellular automata?

   b) Write the Grow-Cut segmentation algorithm for cellular automata. [2+8=10]

6. a) What are the advantages of JPEG2000 over JPEG?

   b) Explain the Embedded Block Truncation Coding (EBTC) used in encoding JPEG2000. [2+8=10]

7. Describe the thresholding method proposed by N. Otsu. How do you extend this method to obtain multiple thresholds? [7+3=10]

8. State and prove the correlation theorem. Show that the Fourier transform of the autocorrelation function of $f(x)$ is its power spectrum. [7+3=10]

9. Consider the following digital signal: [1, 2, 4, 1, -1, -2, -1, 1]. Construct the tree wavelet expansion of this signal using the following wavelet filter: [0.1294, 0.2241, -0.8365, 0.4830]. [10]

10. (a) Write short notes on image registration.

    (b) Define mutual information. How do you compute mutual information between two images. [5+(2+3)=10]

11. Consider the following block of gray levels:

0	0	1	1
0	0	1	1
0	2	2	2
2	2	3	3

    Construct the gray level co-occurrence matrices for angle $\theta = 0°, 45°, 90°, 135°$, considering unit pixel distance, and compute the angular second moment for each case. [(4x2)+2=10]

12. Compute the time dispersion and spectral bandwidth of the following Gaussian signal: $f(t) = e^{-\frac{t^2}{2\sigma^2}}$. Prove that the signal $f(t)$ achieves the minimum of the uncertainty inequality. You may use the following two results:

$$(i) \int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}; \text{ and } (ii) \int_{-\infty}^{\infty} x^2 e^{-x^2} dx = \frac{\sqrt{\pi}}{2}.$$

    [(4+4)+2=10]

INDIAN STATISTICAL INSTITUTE

End Semestral Examination

M. Tech (CS) - II Year, 2012-2013 (Semester - IV)

Topics in Algorithms and Complexity

Date : 24.04.2013 Maximum Marks : 100 Duration : 3.5 Hours

Note: Answer as much as you can, but the maximum you can score is 100.

(Q1) Consider the following algorithm for sorting. Given an input array $\mathcal{A}$ of n distinct numbers, we keep shuffling $\mathcal{A}$ randomly until it is sorted. The shuffling should generate any permutation of the input independently and equally likely.

(a) Design such a shuffling algorithm and argue why your algorithm returns any permutation independently and equally likely.

(b) Analyze the expected time complexity of the said sorting algorithm.

$[5 + 5 = 10]$

(Q2) **(a)** You are given a list of n candidates whom you interview one per day. The recruitment is done according to the following strategy.

Randomly permute the list of candidates and interview the candidates according to this order. Let a be the best candidate among the first $i - 1$ candidates. Interview a candidate i. If i is better than a, then fire a and hire i; else retain a. The cost of hiring a candidate is C_h and the cost of interviewing a candidate is C_i; assume C_i is much less than C_h.

Analyze the above randomized algorithm.

(b) Now consider an on-line version of the problem where the interview is conducted according to the following policy.

Candidates are interviewed one at a time and after the interview, we either tell her that she is not wanted or hire her and stop interviewing any further, i.e. remaining candidates are not interviewed any more. The goal of the problem is to maximize the probability that we get the best person assuming that they come in random order.

The algorithm to solve the problem proceeds in the following fashion. Randomly permute the candidate list. Fix a positive integer $k < n$. Interview the first k candidates and reject them all. Continue to interview the rest of the candidates; hire the first candidate that is better than all preceding candidates. If it turns out that the best candidate was among the first k interviewed, then we hire the n-th candidate.

Note that the bigger k is, the larger the chance that we see the best candidate in the first k and we do not hire the best one. Also, the smaller k is, the larger the chance that we stop too soon.

Determine for each possible value of k, the probability that we get the best candidate and then find the best possible k to implement the strategy.

$[5 + 10 = 15]$

(Q3) **(a)** Let $X_1, \ldots, X_n$ be independent random variables with $\Pr(X_i = 1) = \Pr(X_i = -1) = \frac{1}{2}$. Let $X = \sum_{i=1}^{n} X_i$. For any $a > 0$, show that $\Pr(|X| \geq a) \leq 2e^{-a^2/2n}$.

(b) Let $X_1, \ldots, X_n$ be independent random variables with $\Pr(X_i = 1) = \Pr(X_i = 0) = \frac{1}{2}$. Let $X = \sum_{i=1}^{n} X_i$ and $\mu = E[X] = \frac{n}{2}$. For any $a > 0$, show that $\Pr(X \geq \mu + a) \leq e^{-2a^2/n}$.

[9 + 6 = 15]

(Q4) A hypergraph $\mathcal{H} = (V, E)$ is a generalization of a graph. V is a set of n vertices. $E = \cup_{i=1}^{m} S_i$, where $S_i \subseteq V$. Each S_i is termed as a hyperedge. Notice that if all S_is are 2-element subsets of V, then a hypergraph becomes a graph.

The problem is to find *balanced coloring* of a hypergraph. Each vertex in V is to be assigned the color RED or BLUE. The *discrepancy* $\mathcal{D}(i)$ of a hyperedge S_i is defined as $|R_{S_i} - B_{S_i}|$, where R_{S_i} and B_{S_i} denote the number of red and blue vertices in the hyperedge S_i. The discrepancy $\mathcal{D}$ of $\mathcal{H}$ is defined to be $\max_i \mathcal{D}(i)$. The goal is to assign colors to the vertices so that $\mathcal{D}$ is minimized.

If each vertex is colored RED or BLUE with probability $1/2$ independently, then obtain high probability tail bounds on $\mathcal{D}$. You can use Chernoff bounds and the union bound. [12]

(Q5) Consider the Erdős-Rényi random graph model $G(n, p(n))$. Deduce the following results to comment on the phase transition behavior of K_4 in $G(n, p(n))$ by showing the existence of a $p_0(n)$ such that

(a) If $\lim_{n\to\infty} \frac{p(n)}{p_0(n)} \to 0$, then $\Pr[G \in G(n, p(n))$ contains K_4 as a subgraph $] \to 0$.

(b) If $\lim_{n\to\infty} \frac{p(n)}{p_0(n)} \to \infty$, then $\Pr[G \in G(n, p(n)))$ contains K_4 as a subgraph $] \to 1$.

[6 + 10 = 16]

(Q6) Consider a weighted version of the MAX SAT (maximum satisfiability) problem with n Boolean variables $(x_1, \ldots, x_n)$, m clauses $(C_1, \ldots, C_m)$ and a positive weight w_i for each clause C_i, $1 \leq i \leq m$. Each clause has at least one Boolean variable. MAX SAT is a NP-Hard problem.

We design a simple randomized approximation algorithm for MAX SAT in the following way. Set each x_i to be TRUE independently with probability $1/2$.

(a) Deduce the approximation ratio of the above randomized algorithm.

(b) Derandomize the above randomized algorithm to obtain a deterministic approximation algorithm with the same approximation ration. Prove your results.

[8 + 12 = 20]

(Q7) Consider the dynamic setting in a skip list $\mathcal{L}$. A *finger search* from x to y in $\mathcal{L}$ is a query where you start from the node that contains x (instead of the root) and find the node in $\mathcal{L}$ that stores y. Assume that x is always in the bottommost list of $\mathcal{L}$.

Design and analyze a randomized algorithm for finger search in $\mathcal{L}$ that has a high probability tail bound. You would get credit if your bounds involve $\text{rank}(x)$ and $\text{rank}(y)$, where $\text{rank}(x)$ and $\text{rank}(y)$ denote the current ranks of the elements x and y in the sorted order of the dynamic set. [15]

(Q8) Design and analyze a randomized incremental algorithm for computing the convex hull of n points in $\mathbb{R}^2$ without using the conflict structure. [10]

(Q9) **(a)** Let $\mathcal{A} = A_1, \ldots, A_i, \ldots, A_n$ be a set of "bad events" and let $D_i \subseteq \mathcal{A}$ be a set such that A_i is mutually independent of all events not in D_i. If there exists a set of real numbers $x_1, \ldots, x_n \in [0, 1)$ such that $\forall i$ $\Pr[A_i] \leq x_i \Pi_{j \in D_i}(1 - x_j)$, then

$$\Pr[\cap_{i=1}^{n} \overline{A_i}] \geq \Pi_{i=1}^{n}(1 - x_i) > 0.$$

(b) Let $\mathcal{A} = A_1, \ldots, A_i, \ldots, A_n$ be a set of "bad events" and let $D_i \subseteq \mathcal{A}$ be a set such that A_i is mutually independent of all events not in D_i. If $\sum_{j \in D_i} Pr[A_j] \leq \frac{1}{4}$, $\forall i$, then

$$\Pr[\cap_{i=1}^{n} \overline{A_i}] \geq \Pi_{i=1}^{n}(1 - 2\Pr[A_i]) > 0.$$

$[15 + 5 = 20]$

INDIAN STATISTICAL INSTITUTE
Second Semester Examination: 2012-2013
M.Tech (CS) II Year
Advanced Database Theory and Applications

Date: 25.04.2013 Maximum Marks – 50 Duration – 2 Hours.

1. An organization is working on some construction projects in different parts of the country. It maintains a database of its employees and the projects under execution. The entire database is maintained at the head-quarter of the organization in Kolkata. Appropriate fragments are maintained at different worksites. The present worksites are in Delhi, Mumbai, Chennai and Kolkata. Delhi site has only Civil Engineers, Mumbai site has only Electrical Engineers and Chennai and Kolkata sites have both Civil and Electrical Engineers. Relevant relations of the database schema are shown below:

   Employee (e_no, e_name, e_type, salary, location)
   Technical (e_no, specialization, p_no)
   Project (p_no, p_name, budget, st_date, expected_completion_date, e_no, location)

   Usually following three monthly reports are generated at each worksite and sent to the head-quarter.

   1) List of the engineers working at each site along with their salary and specialization, so that their salary can be transferred to the appropriate bank accounts at their respective worksites.
   2) For the same purpose, list of administrative employees at each location is also maintained.
   3) Progress report about the projects getting executed at each worksite after computing the time left for completion and the budgeted amount yet to be spent.

   Indicate the horizontal/vertical fragments of the above relational schema that need to be maintained at each worksite, in order to generate the above reports. Following assumptions may be made:
   a) Only relevant attributes required to generate the above reports are maintained at each worksite.
   b) Attribute 'e_type' in relation Employee can either be technical or administrative.
   c) Attribute 'specialization' in relation Technical indicates type of engineering degree a technical employee may have.
   d) Kolkata being the head-quarter, the entire database is maintained there.

   (13)

2. Two sets of relational schemas are given below:

Set-1	Set-2
R_{11} (a, b, c, d, e)	R_{21} (p, q, r, a)
R_{12} (x, y, z)	R_{22} (m, n, x)

   Further information about the attribute relationships in the two schemas are:
   1) The underlined attributes are the primary keys.
   2) Relations R_{11} and R_{12} in Set-1 are similar to the relations R_{21} and R_{22} respectively in Set-2.
   3) Attribute 'a' in R_{11} and R_{21} are synonyms but attribute 'x' in R_{12} and R_{22} are homonyms.
   4) All other attributes are different from each other.

   Using the above information, generate a global conceptual schema.

   (12)

3. There are three sites S_1, S_2 and S_3. S_1 and S_2 maintain the data for relations R_1 and R_2 respectively. Additionally, S_1 knows the schema of R_2 and S_2 knows the schema of R_1 without having the data of these relations. S_3 has no data or schema related information. Relations R_1 and R_2 are to be joined, and the result is to be sent to S_3. Indicate a method of achieving such join without sending the entire R_1 or R_2 relation to any of the sites who are not maintaining them. Partial transfer of data is however allowed.

(10)

4. Let A be an object belonging to the security class X and B be a subject belonging to the security class Y. Explain in each of the following cases whether B will be permitted to execute the desired operations on A when the concerned DBMS implements both mandatory and discretionary access control mechanisms. The mandatory system follows the Bell-Lapadula model. Discretionary system can be used only if both the subject and object belong to the same security class.

   i) The owner of A has granted explicit read and write privileges to B, and B wants to execute these operations when X>Y.

   ii) Under the same set of discretionary access rights, B wants to read and write on A when X=Y.

   iii) B wants to read and write on A when X<Y, but the owner of A has not granted any privilege to B. B, however, has received the required privileges from another subject C who in-turn has received them from the owner of A with grant option. C belongs to security class X.

(15)

-x-

INDIAN STATISTICAL INSTITUTE

Second Semestral Examination: 2012-2013

M. Tech. (CS) II year

Data Mining and Knowledge Discovery

Date: 27.04.13 Maximum Marks: 80 Duration: 3 hours

1. (i) What is the advantage of density-based clustering?
   (ii) Describe algorithm DBSCAN.
   (iii) In what way does OPTICS improve upon DBSCAN? [4+8+4=16]

2. (i) Draw and explain a typical building block of a feature selection process.
   (ii) Rank the features of the following data set according to ReliefF method. Show your calculations in detail.

F1	F2	F3	F4	Class labels
4.3	3.0	1.1	0.1	1
5.8	4.0	1.2	0.2	1
5.7	4.4	1.5	0.4	1
5.4	3.9	1.3	0.4	1
5.1	3.5	1.4	0.3	1
5.7	3.8	1.7	0.3	1
7.0	3.2	4.7	1.4	2
6.4	3.2	4.5	1.5	2
6.9	3.1	4.9	1.5	2
5.5	2.3	4.0	1.3	2
6.8	3.2	5.9	2.3	3
6.7	3.3	5.7	2.5	3
6.7	3.0	5.2	2.3	3

   (iii) Write a short note on feature clustering. [4+8+4=16]

3. (i) What is the perceptron learning rule? When does it fail?
   (ii) Explain the role of multilayer perceptron towards classification.
   (iii) What are some limitations of backpropagation learning? Elucidate some ways of resolving them. [5+8+5=18]

4. (i) How is decision tree different, as a classifier, with respect to neural network?
   (ii) Explain algorithm SPRINT, pointing out its merits and demerits in the context of data mining. [4+12=16]

5. (i) Define support and confidence in association rules.
   (ii) How does FP-tree help in rule mining? Explain with an example. [5+9=14]

Indian Statistical Institute

Semester Examination 2012-13
M. TECH.(CS) II Year
Subject: Document Processing and Retrieval
Full Marks: Duration: 3 hrs.

27.04.13

(Answer all questions)

1. Generate a feature based tree classifier for the recognition of the following printed alphanumeric characters (I, J, c, e, q, d, u, n, Q, O). [10]
2. What are the differences between holistic recognition and segmentation based recognition? What do you mean by real-word errors and non-word errors? Discuss about different errors that may occur during optical character recognition. [3+2+5]
3. Discuss about OCR technology for multi-script documents. Discuss about two robust features to separate printed Devnagari script from Roman script. What is a profile based feature? How can such a feature be used for touching character segmentation? Mention one of the drawbacks of this feature for touching character segmentation. [4+5+3+5+3]
4. What is Word-spotting? Discuss about a scale invariant word spotting method for Bangla printed words. [3+7]
5. a) Consider all possible files of size n-bits and we want to compress them to files of size n/2-bits. How many of the n-bit source files can at most be compressed? [5]

   b) Define an m-th order Golumb code. Find the codewords of 4-th order Golumb code for ten symbols expressed as 0, 1, 2 ... 9. How to choose m for efficient coding of pixel run lengths? [4 + 8 + 3]
6. For seven symbols of probabilities 0.25, 0.20, 0.15, 0.15, 0.10, 0.10 and 0.05, construct a Shannon- Fano code and a Huffman code. Compare the efficiency between them. [8+2]
7. a) Describe Bagging, boosting and random forest approaches in connection with classifier combination. [3 + 3 + 4]

   (b) Show that in the approach of normalizing the scores of individual classifiers using *Informational confidence*, the performance function $p(K_i)$ for i-th classifier confidence K_i is given by the exponentially distributed confidence values

   $$p(K_i) = 1 - \exp(-K_i/E)$$

   where E is a multiplying factor influencing the scale.

   Using this relation describe an approach of performance estimation. [5 + 5]

INDIAN STATISTICAL INSTITUTE
M. Tech (CS) II year : 2012-2013
Quantum Information Processing & Quantum Computation
Semestral Examination

Date: 29. 04. 2013 Maximum Marks: 100 Time: 3 Hours
Answer any 5 questions.

1. (a) What is an entangled state?
   (b) Explain how a Hadamard gate and a CNOT gate can be used to obtain a two-qubit entangled state.
   (c) Describe how quantum entanglement is used in teleportation.

   4+6+10 = 20

2. (a) What is a Dicke State? Give an example considering 4-qubits.
   (b) Write an algorithm to efficiently construct an n-qubit Dicke State.

   5+15 = 20

3. (a) Outline the steps of BB84 Quantum Key Distribution Protocol.
   (b) Briefly explain the CNOT attack on this protocol.

   10+10 = 20

4. (a) What is Quantum Fourier Transform (QFT)?
   (b) Describe the QFT steps with the details of the quantum circuits involved. [You may explain using three qubits.]

   5+15 = 20

5. (a) Explain how order finding is related to factorization.
   (b) Briefly explain an overview of the steps involved in Shor's quantum algorithm for factorization.

   10+10 = 20

6. (a) Describe the three qubit bit flip code (error correcting code) with complete details of the quantum circuit involved.
   (b) Give an example that it cannot correct more than one error.

   15+5 = 20

7. (a) Describe the Deutsch-Jozsa algorithm.
   (b) Consider that a Boolean function $f(x_1, x_2, x_3) = x_1 \oplus x_2 x_3$ is used in the Deutsch-Jozsa algorithm as the unitary transform U_f. Explain the outcome after the measurement.

   10+10 = 20

8. Write short notes on (any two): (a) Grover algorithm, (b) Quantum Phase Estimation, (c) Remote State preparation, (d) Super Dense Coding.

   10+10 = 20

INDIAN STATISTICAL INSTITUTE

Semestral Examination: (2012 – 2013)

M.Tech. (CS) II Year

Parallel Processing: Architectures and Algorithms

Date: April 29, 2013 Total Marks: 116 Duration: 3 hours

NOTE: You may answer all questions but the maximum marks attainable is 100.

1. a) Show the schematic diagram of a Parallel Random Access Machine (PRAM), and mention the underlying assumptions of the model. Compare the PRAM model with SIMD, MIMD and CUDA GPU architectures and mention which architecture can best model the PRAM machine.
b) For a given program, 80% of the code is executable simultaneously by 20 processors. The rest of the code is to be executed sequentially. Find the maximum speed-up achievable using i) Amdahl's law and ii) Gustafson's law respectively. Which bound is more achievable if the workload is scalable?
[(4+6)+(4+4+2)=20]

2. a) Prove that an *(n-1)*-node complete binary tree is not a subgraph of an *n*-node hypercube for $n=2^k$, $k>2$.
b) Given a simple polygon P and a stream of query points Q, describe a parallel algorithm with optimal *AT-cost* to find if Q is within P. Mention the model of computation and the interconnection network used.
[6+ 12= 18]

3. a) Find the diameter of the shuffle-exchange interconnection network.
b) Develop an *O(log n)* parallel algorithm for transposing an $n \times n$ matrix, where n is a power of 2. Assume that the processors are interconnected by a perfect shuffle network. Show that the *AT-cost* of the system is sub-optimal.
[4+(7+4) = 15]

4. Given a weighted graph G(V,E), describe a cost-optimal parallel algorithm for finding the minimum spanning tree of G(V, E). Use a suitable shared memory model of SIMD computer. Analyze the time complexity of the algorithm to show that it is cost-optimal.
[10+5= 15]

5. Explain how the Newton's method for solving non-linear equations can be implemented efficiently on a shared memory parallel processor to solve the equation $f(x) = 0$. Assume that the equation $f(x) = 0$ has one and only one root in an interval (a, b). Which of the models SIMD or MIMD will be better and why?
[12+3=15]

6. Draw the customized interconnection network to interconnect the elementary processors capable of *compare and exchange* operations only to implement the Batcher's odd-even merge sort algorithm for sorting a sequence of 8 elements. Describe the algorithm and prove its correctness.
[4+6+5 =15]

7. Answer in brief:
(a) Find the minimum number of middle-stage switches in a 3-stage Clos network to make it non- blocking. Justify your answer.
(b) Prove that *(2n-3)* is an upper bound on the diameter of an *n*-pancake graph.
(c) Prove that if each node of an N×N mesh contains one packet to be routed to a unique destination following the farthest-first strategy, routing of all the packets can be completed in (2N-2) steps only.
[3 × 6 =18]

INDIAN STATISTICAL INSTITUTE

Final Examination:(2012-2013)

MTech C.S. 2nd Year

Advanced Digital Signal Processing

Date: 29.04.2013 Maximum Marks: 100 Duration: 3 hours

Note: The marks add up to 116. The maximum you can score is 100. The exam is open-book, open-notes. Use of calculators is permitted.

1. A wide-sense stationary random process $x[n]$ has zero mean and autocorrelation function $0.6^{|n|}$.

   (a) Determine a causal and stable system for generating $x[n]$ from white noise.

   (b) Determine the whitening filter for $x[n]$.

   (c) Let $x[n]$ be the input for the linear system represented by the difference equation

   $$y[n] = 0.8y[n-1] + x[n] - x[n-1]$$

   Determine the power density spectrum of the output.

   (d) Determine the autocorrelation of the output.

   [8+5+8+5]

2. An AR(1) process satisfies the difference equation

   $$s[n] = 0.8s[n-1] + v[n]$$

   where $v[n]$ is a white noise sequence with variance 0.36. Consider a signal

   $$x[n] = s[n] + w[n]$$

   where $w[n]$ is a white noise sequence with variance 0.49

   (a) Determine the autocorrelation sequences $\gamma_{ss}[m]$ and γ_{xx}.

   (b) Design a Wiener filter of length 2 to estimate $s[n]$ and determine the corresponding minimum mean square error.

   [(8+5)+(8+4)]

3. Derive the parameters of the lattice-ladder structure implementing the system represented by

$$H(z) = \frac{(1-0.4z^{-1})(1-0.5z^{-1})}{(1-0.3z^{-1})(1-0.7z^{-1})}$$

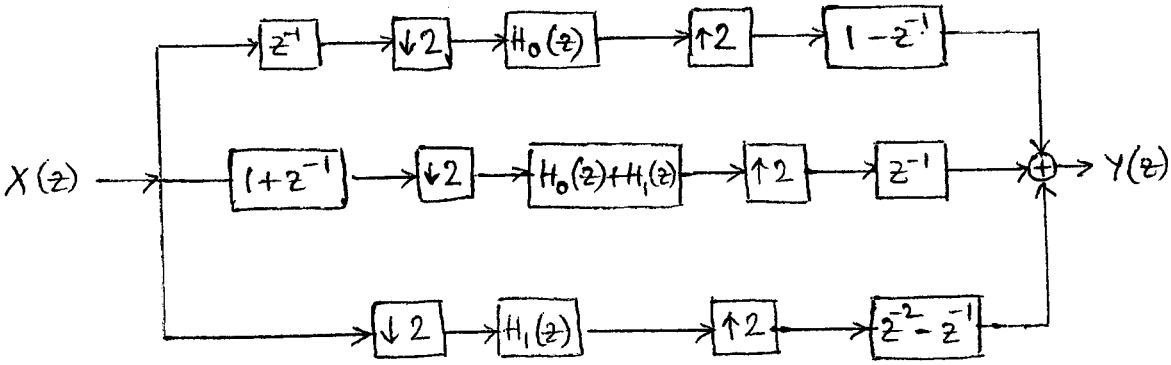
[10+10]

4. Let $H(z)$ be a causal FIR filter of length N, where N is even, have a 2-band polyphase decomposition $H(z) = E_0(z^2) + z^{-1}E_1(z^2)$. Determine $H_0(z^2)$ and $H_1(z^2)$ in terms of $E_0(z^2)$ and $E_1(z^2)$, such that

$$H(z) = (1+z^{-1})H_0(z^2) + (1-z^{-1})H_1(z^2)$$

[8]

5. Determine the overall transfer function $Y(z)/X(z)$ for the system shown below. Is it alias-free? [12+5]



6. A bandlimited continuous time signal has a bandlimited power spectrum that is zero for $|\Omega| \geq 2\pi(10^4)$ rad/s. The signal is sampled at a rate of 20000 Hz for a duration of 10s. For power spectral estimation by the method of periodogram averaging,
Determine

(a) the total number of samples and the minimum length N if the power spectrum estimates must be obtained at equally spaced frequencies no more than 10 Hz apart. Assume that radix 2 FFTs are used for the periodogram computations.

(b) give a method of reducing the variance of the spectrum estimates by a factor of 10 for the same frequency spacing.

[5+5=10]

7. Show that the symmetric Bartlett window defined as

$$\begin{cases} 1 - |m|/M & \text{if } |m| \leq M-1, \\ 0 & \text{otherwise,} \end{cases}$$

is the aperiodic autocorrelation of the sequence $u[n] - u[n-M]$. [10]

INDIAN STATISTICAL INSTITUTE

Semestral Examination: (2012 - 2013)

Course Name: M.Tech (CS) **Year: 2nd year**

Subject Name: Neural Networks & Applications

Date: April 29, 2013 **Maximum Marks: 100** **Duration: 3 hrs**

Answer all the questions.

1. Show how the Oja's model of Principal Component Analysis Network extracts the first and second principal components from a set of n-dimensional samples. State clearly all the assumptions made in this regard. [25]

2. Consider a pattern classification problem where n-dimensional input patterns are distributed in l classes. These patterns need to be classified using a multilayer perceptron (MLP) with one hidden layer consisting of m nodes. Derive the learning rule for weight updation under the gradient-descent technique for minimization of error. Assume sigmoid functions as activation functions of all the hidden and output nodes, along with necessary bias. [25]

3. a) Consider the patterns (1,1,1) and (-1,-1,-1). Show how they can be stored in a Hopfield network.

b) Assume that the activation of a node of the above Hopfield network at its stable state is either $+1$ or -1. Which of the above fundamental memories will be retrieved in response to the following input patterns: (1,1,-1), (1,-1,-1), (1,-1,1), (-1,-1,1), (-1,1,1) and (-1,1,-1).

c) Derive an expression for the storage capacity of a Hopfield network when most of the fundamental memories are retrieved correctly. [10 + 15 + 15 = 40]

4. a) Consider a ten-dimensional data set containing 100 points. The data is known to have 4 groups. Design a Self Organizing Feature Map for clustering this data. Explain its learning rule.

b) Why is the learning rule called competitive learning? [(6+2) + 2 = 10]

INDIAN STATISTICAL INSTITUTE

Semestral Examinations

M.Tech(CS)-II Year, 2012-2013 (Semester-IV)

Distributed Computing Systems

Date: April 29, 2013 Maximum Marks: 100 Duration: **3 Hours**

Note: The question paper carries a total of 115 marks. You can answer as much as you can, but the maximum you can score is 100.

1. (a) Explain the types of failures (such as deadlock, starvation and others) that can happen under the following situations:
   i. The channels do not satisfy FIFO delivery in Lamport's mutual exclusion algorithm.
   ii. A control message is lost in Huang's termination detection algorithm
   iii. The channels do not satisfy FIFO delivery in Chandy-Lamport's algorithm for Global state recording.

   (b) Give an example to show that in the AND request model, false deadlocks can occur due to deadlock resolution in the distributed system.

   (c) Describe the Ricart-Agarwala algorithm for mutual exclusion. Prove the correctness of the algorithm.

   ((2+2+4)+5+(7+5)=25)

2. (a) Describe the Dijkstra's self stabilizing token ring algorithm. Prove the correctness and convergence of the algorithm.

   (b) Describe Chandy-Lamport's Global state recording algorithm.

   ((4+8)+10=22)

3. (a) Is deterministic leader election possible in a synchronous ring in which all but one processors have the same identifier? Either give an algorithm or prove an impossibility result.

   (b) Design a *leader-election* algorithm for asynchronous ring model. Prove the correctness of your algorithm. Establish the time and message complexities of your algorithm.

   (5+(8+3+4)=20)

4. (a) Design an efficient distributed algorithm to color a tree with three colors. What is the time complexity of the algorithm?

(b) Explain how a solution to the Byzantine agreement problem can be used for a solution to the problem of consensus and interactive consistency.

(c) Prove that a Byzantine agreement cannot be reached among three processors, where one processor is faulty.

((9+1)+8+5=23)

5. (a) State and prove the 0-1 lemma for sorting algorithms.

(b) Prove or disprove the following statements:

i. Given any correct sorting network, adding another comparator at the end destroys the sorting property.

ii. Every correct sorting network needs to have at least one comparator between each two consecutive wires.

iii. A network which contains all $\binom{n}{2}$ comparators between any two of the n wires, in whatever order they are placed, is a correct sorting network.

(c) Prove that any sorting network on n inputs has depth at least $log_2(n)$.

((2+6)+(4+6+4)+3=25)

INDIAN STATISTICAL INSTITUTE

End Semester Examination

M. Tech (CS) - II Year, 2012-2013 (Semester - IV)

Computational Complexity

Date : 29.04.2013 Maximum Marks : 100 Duration : 3 hours

Answer any five questions.

(Q1) Discuss briefly the relationship, if any, between each of the following pairs of complexity classes:

(a) DTIME(2^n) and DTIME(3^n)

(b) NSPACE(2^n) and DSPACE(5^n)

(c) DSPACE(n) and DTIME($(\log n)^n$)

(d) BPP and P/poly.

[4*5=20]

(Q2) (a) Show that if P = NP then P = PH.

(b) Recall that NP^{SAT} is the class of languages that are recognized by non-deterministic polynomial time Turing machines with an oracle for the satisfiability problem. Show that $\mathsf{NP}^{SAT} = \Sigma_2^P$.

(c) Argue whether there exists a complete problem for the polynomial hierarchy.

[6+10+4=20]

(Q3) (a) Show that $\{G|G$ is strongly connected$\} \in$ NL.

(b) Prove that TISP($n, \log n$) $\neq$ TIME(n) $\cup$ SPACE($\log n$).

[10+10=20]

(Q4) (a) Show that the transitive closure of a $n \times n$ Boolean matrix can be computed by a circuit of depth $O(\log^2 n)$.

(b) Show that the binary addition function is in AC^0.

(c) Prove that NC $\subseteq$ P.

[8+5+7=20]

(Q5) (a) Let M be a probabilistic polynomial time Turing Macine and let A be a language where, for some fixed $0 < \epsilon_1 < \epsilon_2 < 1$,
$w \notin A$ implies $Pr[M \text{ accepts } w] \leq \epsilon_1$, and
$w \in A$ implies $Pr[M \text{ accepts } w] \geq \epsilon_2$. Show that $A \in \mathsf{BPP}$.

(b) Prove that $\mathsf{dIP} = \mathsf{NP}$. Comment on the power of a probabilistic verifier.

[10+(7+3)=20]

(Q6) (a) Give a 2-factor approximation algorithm for *MAX-3SAT* problem.

(b) What does a $\mathsf{PCP}(r(n), q(n))$ verifier signify?

(c) Prove that $\mathsf{PCP}(\log n, 1) \subseteq \mathsf{NP}$. Comment on whether the converse is true.

[6+5+(6+3)=20]

INDIAN STATISTICAL INSTITUTE

M. Tech. (CS) II Year (2012-13), II semester

Semestral Examination

ADVANCED PATTERN RECOGNITION

Date: 01.04.13 Duration: 195 minutes Marks: 100

Note: Answer all questions.

1. (a) Describe k-nearest neighbor based data condensation algorithm.
   (b) Describe a dissimilarity measure between two features and state a feature selection algorithm using it. [4+7=11]

2. Suppose you have two 2-dimensional normal populations $N(\mu_1,\Sigma)$ and $N(\mu_2,\Sigma)$ where $\mu_1' = (0.0,0.0), \mu_2' = (1.0,2.0), \Sigma = \begin{pmatrix} 1 & -0.7 \\ -0.7 & 2 \end{pmatrix}$. Let the prior probabilities of the populations be 0.5 and 0.5.

   (a) Find the Mahalanobis distance between the two populations.
   (b) Find the Bayes decision rule for separating the two populations. Also find its probability of misclassification in terms of standard normal probabilities. [5+(3+7)=15]

3. (a) Describe crossover and mutation operations in Genetic Algorithms.
   (b) Describe a way of tackling the problem of clustering using genetic algorithms. [(7+4)+6=17]

4. Let $x_1' = (0.0,0.0), x_2' = (1.0,0.0), x_3' = (0.0,1.0), x_4' = (1.0,1.0)$ and $x_5' = (2.0,2.0)$. Let there be two classes and let the first two points belong to the class 1, and the rest of them belong to class 2. Let $x + y = 0.5$ be the starting straight line for Perceptron algorithm. Let $\lambda = 0.1$ be the given learning rate. Apply Perceptron algorithm and find the discriminating straight line between the classes. [15]
5. Describe Fuzzy C means algorithm. [15]
6. (a) Describe the basic steps of gradient descent optimization algorithm.
   (b) Let $f(x,y) = 4x^2 - 12xy + 9y^2 - 10$ where x and y are real numbers. Minimize f by using gradient descent technique. [3+12=15]
7. Write short notes on the following.
   (a) k-fold cross validation
   (b) VC dimension
   (c) Support vectors

   [4+4+4=12]

INDIAN STATISTICAL INSTITUTE

Second Semestral Examination (2013)

M.Tech. (Computer Science) Second Year

Natural Language Processing

Date: 03.05.2013 **Time: 2 h 30 min** **Maximum Marks: 50**

Answer all questions.

Question 1.

(a) Consider a simple Probabilistic Context Free Grammar (PCFG) as given below. The non-terminals are S, NP, PP, VP, P, and V. The start symbol is denoted by S. The terminals are the words in italics.

S →	NP VP	1.0	NP →	NP PP	0.4
PP →	P NP	1.0	NP →	*astronomers*	0.1
VP →	V NP	0.7	NP →	*ears*	0.18
VP →	VP PP	0.3	N →	*saw*	0.04
P →	*with*	1.0	NP →	*stars*	0.18
V →	*saw*	1.0	NP →	*telescopes*	0.1

Using the above grammar, there are two different parses of the sentence *astronomers saw stars with ears* are possible. Show these two parses and compute the probabilities of these two parses. Compute the sentence probability.

[(2 + 2) + (2+2) + 2 = 10]

(b) The following table shows the calculation of inside probabilities $[\beta_i(p, q)]$ for the sentence *astronomers saw stars with ears* for the simple grammar given above in (a). The recursive computation of inside probabilities is done starting along the diagonal, and then moving towards the top right corner.

	1	2	3	4	5
1	$\beta_{NP} = 0.1$		$\beta_S = ?$		$\beta_S = ?$
2		$\beta_{NP} = 0.04$ $\beta_V = 1.0$	$\beta_{VP} = 0.126$		$\beta_{VP} = ?$
3			$\beta_{NP} = 0.18$		$\beta_{NP} = ?$
4				$\beta_P = 1.0$	$\beta_{PP} = 0.18$
5					$\beta_{NP} = 0.18$
	astronomers	*saw*	*stars*	*with*	*ears*

Complete the above table by computing the entries marked by "?". [4 x 2 = 8]

[**Hint**: The value for the cell (2, 3) is calculated as: $P(VP \rightarrow V\ NP)\ \beta_V(2, 2)\ \beta_{NP}(3, 3) = 0.126$]

(c) What value would you get for the cell (2, 5) if the bracketed corpus notation is used and the sentence is given in the following form: (*astronomers*(*saw*(*stars*(*with ears*))))?

[4]

Question 2.

Consider the following discourse:

Ram took Sita for a vacation to Lanka but she did not like the place. It was raining heavily and they forgot to bring an umbrella. They realized how useful an item it was.

Assume you have an anaphora resolution system (ARS) and when the above discourse is input to the system, the following result is generated (each anaphoric word is tagged with the word number in the discourse and for an anaphoric phrase the number of the first word of the phrase in the discourse is given):

she (10)	→	Sita
the place (14)	→	not retrieved
it (16)	→	Lanka
they (21)	→	Lanka, Sita
they (27)	→	umbrella, place
it (33)	→	umbrella

Evaluate the above result by using the MUC metric.

[8]

Question 3.

Assume there are four languages (L_1, L_2, L_3, and L_4) and you are given a corpus (sample text) for each language. Formulate a method to identify the language (one out of the given four) in which a short segment of text is written. For example, each of the following lines is text in a different language:

doen is ondubbelzinnig uit
pretender a un emploi
uscirono fuori solo alcune
look into any little problem

Your method will take one line as input and identify the language in which the line is written. Explain your answer by providing pseudocode of your proposed method.

[10]

[Hint: This is a classification task, in which you may use some of the language modeling techniques taught in your class.]

Question 4.

Consider a Bengali to Hindi machine translation system that because of NLP resource constraints basically does word wise translation. Here is an example:

Bengali: ফুলে ফুলে বাগান ভরে গেছে। (/phule /phule /bAgAn /bhore /gechhe)
Word wise translation: ফুলে/phule → फूल मे; ফুলে/phule → फूल मे; বাগান/bAgAn → बगीचा; ভরে/bhore → भरके; গেছে/gechhe → गया ; So the translation in Hindi becomes: फूल मे फूल मे बगीचा भरके गया। However, the actual translation should have been: फूलों से बगीचा भर गया।

What minimum additional NLP resource(s) would you require to produce the correct translation for the above example? Explain your answer. [10]

Indian Statistical Institute

M. TECH. (CS) 2 Year : 2012–2013
Semester Examination
Subject: Information and Coding Theory
Date: 03/05/2013 Time: 3 hours Marks:100

Note: Notations used are as explained in the class.

1. Prove that for a binary symmetric channel with crossover probability $p < 1/2$, the maximum likelihood decoding rule is the same as the nearest neighbour decoding rule. [10]

2. Prove that any linear code C is *equivalent* to a linear code C' with a generator matrix in the *standard* form. [8]

3. Suppose the parity check matrix of a $[n, k]$ linear code is $[A|I_{n-k}]$. Show that its generator matrix is $[I_k| - A^T]$. [10]

4. Prove that a code with minimum distance d can correct $\lfloor \frac{d-1}{2} \rfloor$ errors. [8]

5. State and prove the Gilbert-Varshamov bound. [10]

6. Construct a single-error correcting linear code with $n = 10$. Exhibit two codewords. [10]

7. Describe a systematic encoding of cyclic codes. Give a non-systematic encoder of RS codes. Describe the majority logic decoding of RS codes. [12]

8. Describe briefly the different stages of decoding BCH codes. [12]

9. Prove that the dual of an MDS code is an MDS code. Also prove that, the dual of a BCH code may not be a BCH code. [8]

10. Let C be an $[n, k, d]$ code over $GF(q)$. Prove that the following statements are equivalent.

    (a) C is MDS.

    (b) Every k columns of a generator matrix G are linearly independent.

    (c) Every $n - k$ columns of a parity check matrix H are linearly independent. [12]

INDIAN STATISTICAL INSTITUTE
End-Semester Examination: 2012-13
Course Name: M.Tech. In Computer Science
Subject Name: Fault Tolerant Computing

Date: 06. 05. 2013 *Maximum Marks: 70* *Duration: 3 hours*

Answer question 1 and any three from the rest.

---

1. For each of the following statements, indicate whether the statement is true or false. For the ones that are true, provide a brief justification. For the ones that are false, provide a small counter-example.
   (a) All LTL properties can be expressed in CTL.
   (b) In the *counter-example guided abstraction refinement* (CEGAR) approach, the abstract machine has fewer runs as compared to the original machine.
   (c) The complexity of both LTL model checking and CTL model checking are linear in the size of the Kripke structure.
   (d) A bounded model checking tool needs to add clauses for loop detection while searching for counter-example traces of LTL properties of the form $\varphi U \psi$.
   (e) Any CNF formula in which there is not more than one positive literal in each clause, is solvable in polynomial time.
   (f) There exist languages, which are accepted by some non-deterministic Buchi automaton but not by any deterministic Buchi-automaton.
   (g) The worst case space complexity of a DPLL SAT solver without learning is polynomial in the size of the original problem.
   (h) The size of the ROBDD of $(a1 \wedge b1) \oplus (a2 \wedge b2)$ for the variable ordering $(a1 < a2 < b1 < b2)$ is less than the one generated for the same function for the ordering $(a1 < b1 < a2 < b2)$.

   [2 X 8 = 16 marks]

2. [Equivalence Checking]
   (a) State the register correspondence problem for equivalence checking between sequential circuits.
   (b) Give an algorithm for solving the register correspondence problem and illustrate its working with an example.
   (c) If no register correspondence exists, can we conclude that the sequential circuits are not equivalent? If not, give a counter-example.

   [4 + 9 + 5 = 18 marks]

P.T.O

3. [Model Checking]

(a) Construct the tableau for the LTL property $\varphi = \neg (pUq)$. Clearly explain the steps.

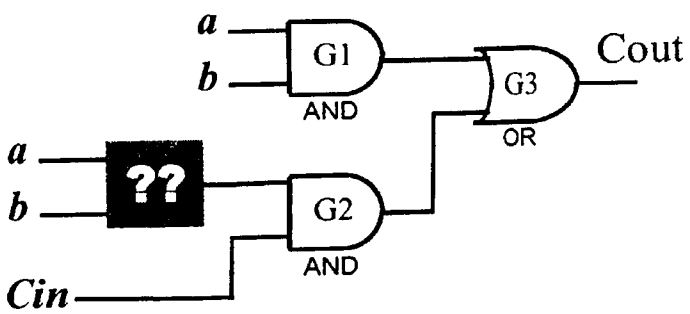
(b) Show a run in this tableau that starts in a state in Sat(φ), but is not an accepting run for φ. How can we capture the accepting runs in this tableau in terms of a Buchi acceptance condition?

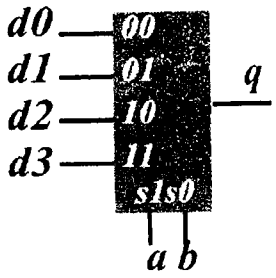
(c) Draw a single Kripke structure that satisfies the CTL property AGAF(p) and the LTL property F$\neg$p.

[8 + 4 + 6= 18 marks]

4. [Network Repair, Unate Recursive Paradigm]

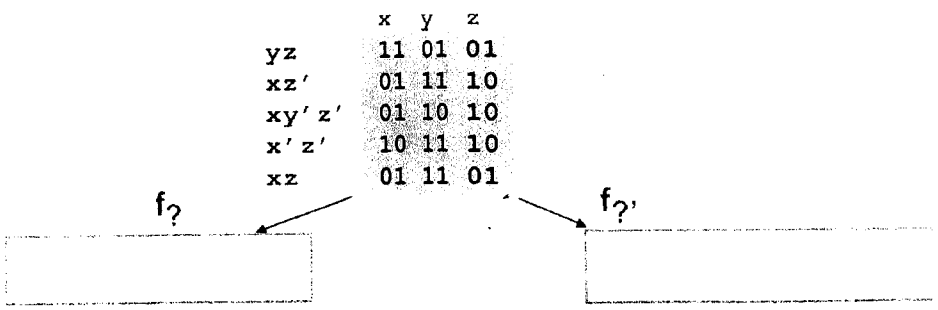
(a) The carry output of a 1-bit adder has the Boolean equation $C_{out} = a \cdot b + (a+b) \cdot C_{in}$, where a and b are the 1-bit numbers we want to add, and C_{in} is the input carry bit. Suppose we know that we have implemented the gate-level logic incorrectly, the gate with the "??" label is the one we suspect is incorrect.





Use the logic network repair procedure and replace the "??" gate with the 4:1 multiplexor shown above, with inputs a, b connected to the s1, s0 select lines respectively. (So, for example, s1s0=10 means q=d2 for this mux.) Use the quantification procedure, and do the Boolean algebra to see what the result tells you about how to fix it.

(b) Consider an intermediate step in our URP tautology recursion. Suppose we have these cubes at one node, and we need to decide on the splitting variable to use to cofactor and recurse. Which variable do we pick, and why?



[10 + 8 = 18 marks]

5. [Theory of Satisfiability]

(a) Consider a clause $(x_1 \vee x_2 \ldots \vee x_n)$, $n > 2$, in a non-CNF formula. How many auxiliary variables are needed for encoding it with Tseitin's encoding? Suggest an alternative way to encode it, using a single auxiliary variable. How many clauses are needed?

(b) Show a formulation in propositional logic of the following problem: given a directed graph, does it contain a Hamiltonian cycle? [**Note**: A Hamiltonian cycle in a graph is a closed path that visits each node, other than the first, exactly once].

[(3 + 3 + 3) + 9 = 18 marks]